scientific reports

Check for updates

OPEN

Human-generative AI collaboration enhances task performance but undermines human's intrinsic motivation

Suqing Wu, Yukun Liu✉, Mengqi Ruan, Siyu Chen & Xiao-Yun Xie

In a series of four online experimental studies (total *N* = 3,562), we investigated the performance augmentation effect and psychological deprivation effect of human-generative AI (GenAI) collaboration in professional settings. Our findings consistently demonstrated that collaboration with GenAI enhanced immediate task performance. However, this performance augmentation effect did not persist in subsequent tasks performed independently by humans. Importantly, transitioning from collaboration with GenAI to solo work led to an increased sense of control of human workers, and was also accompanied by significant decreases in intrinsic motivation and increases in feelings of boredom. These results highlight the complex dual effects of human-GenAI collaboration: It enhances immediate task performance but can undermine long-term psychological experiences of human workers.

"*But what if the opposite ends up happening, and AI takes on all the fun stuff?*"

- An excerpt from an article titled "The End of Coding as We Know It", written by Aki Ito and appeared in Business Insider on April 26, 2023.

Recent advancements in artificial intelligence (AI) have largely shifted its role from mere task automation to human capabilities augmentation[1]. AI systems were once limited to routine and repetitive tasks, but now they can collaborate with human workers on complex and cognitively demanding tasks[2]. The advent of generative AI (GenAI), capable of producing high-quality content such as text, images, and synthetic data, is reshaping the nature and structure of many professional tasks[3]. Increasingly, professional tasks are becoming a synergy of human and AI contributions, marking a shift towards collaborative human-AI work dynamics[4–6].

While no occupation exists where GenAI can entirely substitute human roles[7], there remain tasks that can be better executed by human workers. The future of employment is more likely to evolve towards a hybrid model, where individuals switch between collaborating with GenAI and working independently[8]. This evolving hybrid work dynamic highlights the need for work redesign and effective task allocation within a job role[9]. In particular, GenAI is increasingly being integrated into tasks that require creativity and problem-solving, complementing human capabilities in producing high-quality outputs[10,11]. Our research focuses on text-based open-ended tasks designed to create professional or creative content, such as composing work emails, drafting performance reviews, and generating ideas for product improvement. These tasks are both cognitively demanding and reflective of the responsibilities commonly encountered in professional workplace settings.

Recent studies have demonstrated that collaboration with GenAI enhances both productivity and the quality of human tasks[4,10–14]. For instance, one study found that GenAI enhanced the productivity of less skilled customer support agents[3], while another showed that programmers using OpenAI tools completed tasks faster than control groups[15]. Beyond efficiency, GenAI has also improved work quality in contexts requiring empathy and creativity. For example, mental health counselors using an AI-enabled chatbot produced more empathetic responses[16], and customer service employees collaborating with AI exhibited greater creative performance[17]. Additionally, ChatGPT has been shown to help professionals produce higher-quality writing with less effort[4]. Collectively, these studies underscore GenAI's potential to enhance diverse aspects of human performance across professional domains.

While immediate performance benefits of collaborating with GenAI are evident, it is also important to examine its *long-term* effects on human workers' psychological experiences and task performance. In occupational settings, tasks that allow for creative freedom and problem-solving are often inherently motivating[18]. However,

School of Management, Zhejiang University, 866 Yuhangtang Road, Hangzhou 310058, People's Republic of China. ✉email: liuyukun@zju.edu.cn

as noted in the opening quote, GenAI may diminish the intrinsically motivating parts that are essential for human's sustained work engagement. For example, creating a performance review enables evaluators to critically analyze an individual's strengths and weaknesses and experience the fulfillment of offering tailored, constructive feedback. When GenAI takes over these aspects, it may reduce the analyzing and crafting processes that make such tasks engaging. Intrinsic motivation refers to the internal drive to engage in activities for people's own sake, driven by personal interest, enjoyment, or the satisfaction derived from the activity itself, rather than external rewards or pressures[19,20]. Given the extensive research indicating the positive impact of GenAI on the immediate task performance, we believe it is necessary to assess its long-term effects on human workers' psychological experiences and long-term performance.

This research aims to explore the long-term implications of human-GenAI collaboration on human workers' psychological experiences and performance. Specifically, we investigate whether the augmentation effect of human-GenAI collaboration extends to the subsequent tasks performed by human workers independently. Additionally, we examine whether such collaboration induces psychological costs, including diminished sense of control, intrinsic motivation, and heightened feelings of boredom. To address these questions, we conducted four experiments, involving texted-based tasks such as report writing, idea brainstorming, and problem-solving. Participants engaged in tasks where GenAI collaboration preceded independent work and compared to those engaging in only solo tasks.

We hypothesize that collaboration with GenAI leads to a spillover augmentation effect, enhancing performance in subsequent human-solo tasks (*RQ1*). Additionally, we aim to investigate and replicate the performance enhancement effect of generative AI on immediate human-GenAI collaboration tasks, as demonstrated in prior research. We argue that the spillover augmentation effect may arise through both cognitive and motivational pathways. From a cognitive perspective, interacting with GenAI introduces diverse perspectives that can stimulate creative thinking, much like interactions with other humans[21–23]. Exposure to novel ideas may encourage a mindset of creativity that persists in subsequent tasks. Additionally, task switching, facilitated by completing a GenAI-assisted task, may also reduce cognitive fixation and promote greater flexibility in problem-solving[24]. Second, GenAI can save human workers' effort[4], which may help individuals feel less fatigued, enabling better engagement with subsequent cognitively demanding tasks. Motivationally, AI assistance can bolster individuals' perceptions of competence[25,26], potentially increasing motivation and effort in subsequent tasks.

Despite the potential benefits of GenAI on task performance, its influence on psychological experiences can be examined through three key aspects identified by Self-Determination Theory (SDT): intrinsic motivation, sense of control, and boredom. Intrinsic motivation, defined as the inherent desire to seek out novelty and challenges[19,20], promotes sustained engagement in activities[27]. According to SDT, intrinsic motivation arises from the fulfillment of three innate psychological needs: autonomy, relatedness, and competence[19,28]. Emerging technologies, such as GenAI, can impact these motivational factors either positively or negatively[7]. Research found that participants collaborating with ChatGPT enjoyed the tasks more[4]. When transitioning to solo work, this contrast may make the task feel less enjoyable for humans, thereby reducing intrinsic motivation. Additionally, AI capabilities can reduce the need for human effort, which in turn lowers human engagement[29]. Therefore, we hypothesize that collaboration with GenAI diminishes intrinsic motivation during subsequent human-solo tasks (*RQ2*).

Sense of control, a core component of SDT, is crucial for task engagement as it enables individuals to feel they are the primary agents of their actions. In the context of GenAI collaboration, the sense of control—defined as perceiving oneself as the initiator of one's actions[30]—is particularly important. When individuals collaborate with AI, their perceived autonomy may diminish if they feel that AI-driven contributions override their own decision-making[31,32]. For instance, prior research has shown that interactions with AI chatbots can reduce customers' sense of autonomy[33]. Given the role of autonomy in maintaining task engagement, we examine sense of control (*RQ3*) to understand whether GenAI collaboration undermines this fundamental psychological need compared to human-solo tasks.

Boredom is another critical variable. Although intrinsic motivation drives engagement with activities by fostering a desire for novelty and challenges, it may not be sufficient to sustain engagement in tasks that lack these elements. Tasks that are less stimulating, especially when compared to more engaging tasks, can seem more mundane and contribute to heightened feelings of boredom[34]. Boredom, characterized by a lack of interest and difficulty focusing[35], is a disruptive state that erodes individuals' well-being[36–38]. Boredom often arises when tasks fail to provide adequate challenges or meaningful engagement[39]. In our context, transitioning from engaging AI-assisted tasks to comparatively mundane solo tasks may heighten perceptions of boredom, making these subsequent tasks less enjoyable[34]. Therefore, we hypothesize that collaboration with GenAI intensifies the feeling of boredom during solo tasks, compared to consistent human-solo work (*RQ4*).

Overview

We conducted four online experimental studies to test our predictions (see Fig. 1 for an overview of Studies 1–3). All studies were designed using Qualtrics and incorporated ChatGPT as part of the experimental setup. Participants were recruited from Prolific. In Studies 1–3, participants were randomly assigned to one of two conditions. In the first condition, labeled as "Collab-Solo", participants collaborated with ChatGPT to complete Task 1 and then completed Task 2 on their own, representing a transition from human-GenAI collaboration to human solo effort. In the second condition, labeled as "Solo-Solo", participants completed both Task 1 and Task 2 independently, without any assistance from GenAI. After each task, participants responded to Likert-scale questions assessing their sense of control, intrinsic motivation, and feelings of boredom. Study 4 expanded the experimental design to include four conditions. In addition to the Collab-Solo and Solo-Solo conditions, two new conditions were introduced: "Collab-Collab", in which participants collaborated with ChatGPT for

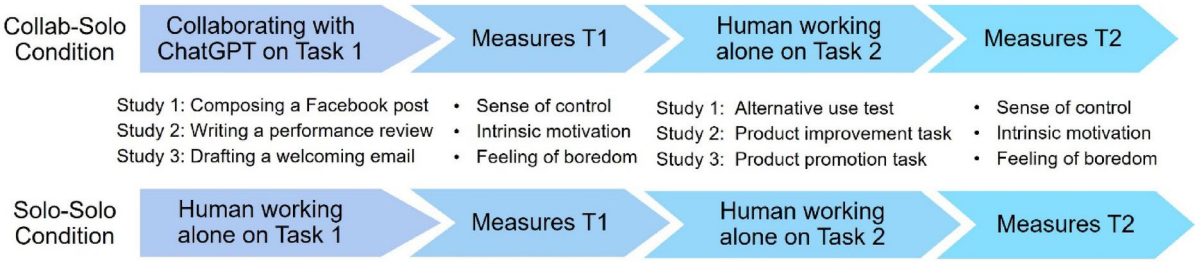


Fig. 1. Overview of Studies 1 to 3. This figure illustrates the experimental setup across Studies 1 to 3. For each study, participants in the Collab-Solo condition (top flow) worked with ChatGPT on Task 1, then proceeded to work alone on Task 2, whereas participants in the Solo-Solo condition (bottom flow) independently completed both Task 1 and Task 2. Measures taken after each task assessed participants' sense of control, intrinsic motivation, and feelings of boredom. Each study adopted different tasks.

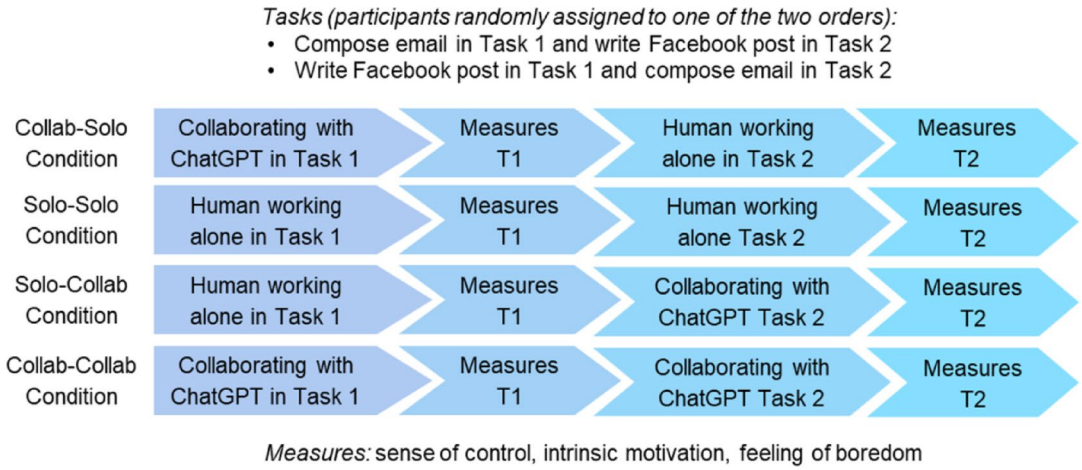


Fig. 2. Study 4 experiment design.

both tasks, and "Solo-Collab", where participants worked independently on Task 1 and then collaborated with ChatGPT on Task 2 (see Fig. 2 for an illustration and the Methods section for a detailed description).

Study 1

In Study 1 ($N = 352$), participants first completed Task 1, which involved drafting a promotional Facebook post with or without ChatGPT's assistance (Collab-Solo vs. Solo-Solo, respectively). After completing Task 1, participants reported their psychological experiences (including sense of control, intrinsic motivation, and feelings of boredom) before proceeding to Task 2, which was an Alternative Uses Test (AUT)[40]. Participants brainstormed innovative uses for a soda can by themselves in the AUT.

Performance augmentation of ChatGPT (RQ1)

We evaluated whether participants collaborating with ChatGPT produced Facebook posts of superior quality on two criteria: how engaging and informative the posts were. Consistent with our hypothesis, results revealed that posts from human-GenAI collaboration received higher scores on quality compared to those from participants working independently ($M_{Collab} = 2.98$, $SD_{Collab} = 0.79$ vs. $M_{Solo} = 2.81$, $SD_{Solo} = 0.72$; $t(350) = 2.15$, $p = .032$, 95% $CI = [0.01, 0.33]$, $d = 0.23$). The small but significant effect size suggests that collaboration with GenAI yields modest improvements in post quality. Specifically, this effect works on producing engaging outputs ($M_{Collab} = 2.89$, $SD_{Collab} = 0.90$ vs. $M_{Solo} = 2.68$, $SD_{Solo} = 0.83$; $t(350) = 2.22$, $p = .027$, 95% $CI = [0.02, 0.39]$, $d = 0.25$), but not on producing informative posts ($M_{Collab} = 3.08$, $SD_{Collab} = 0.79$ vs. $M_{Solo} = 2.94$, $SD_{Solo} = 0.75$; $t(350) = 1.72$, $p = .087$, 95% $CI = [-0.02, 0.30]$, $d = 0.18$). This finding suggests the performance augmentation effect of GenAI. Specifically, the effect sizes (Cohen's d) across these comparisons suggest small but meaningful improvements in the quality of posts generated through human-GenAI collaboration, primarily driven by more engaging rather than more informative content (see Table 1; Fig. 3 for the result summary).

Human-GenAI collaboration and the subsequent human-alone task performance (RQ1)

We further assessed whether participants from the human-GenAI collaboration condition outperformed their counterparts in the subsequent alternative uses test (see Table 2; Fig. 4 for the result summary). In this task, all participants were asked to brainstorm as many alternative uses for a soda can as possible, without the aid of AI.

DV	Engaging		Informative		Overall Quality	
Condition	Collab-Solo	Solo-Solo	Collab-Solo	Solo-Solo	Collab-Solo	Solo-Solo
Mean	2.89	2.68	3.08	2.94	2.98	2.81
SD	0.90	0.83	0.79	0.75	0.79	0.72
t (*df*)	**2.22 (350)**		1.72 (350)		**2.15 (350)**	
p	**0.027**		0.087		**0.032**	
95% *CI*	**[0.02, 0.39]**		[-0.02, 0.30]		**[0.01, 0.33]**	
Cohen's *d*	**0.25**		0.18		**0.23**	

Table 1. Summary of main findings of augmentation effect of collaboration with ChatGPT for Study 1, Task 1 performance. DV = dependent variable, *SD* = standard deviation, *CI* = confidence interval, Cohen's *d* is the standardized effect size. Comparisons were conducted between Collab-Solo (transitioning from collaboration to solo work) and Solo-Solo (working solo in both tasks) conditions. *T*-tests were used to evaluate mean differences for each dependent variable. Significant results are bolded for emphasis.

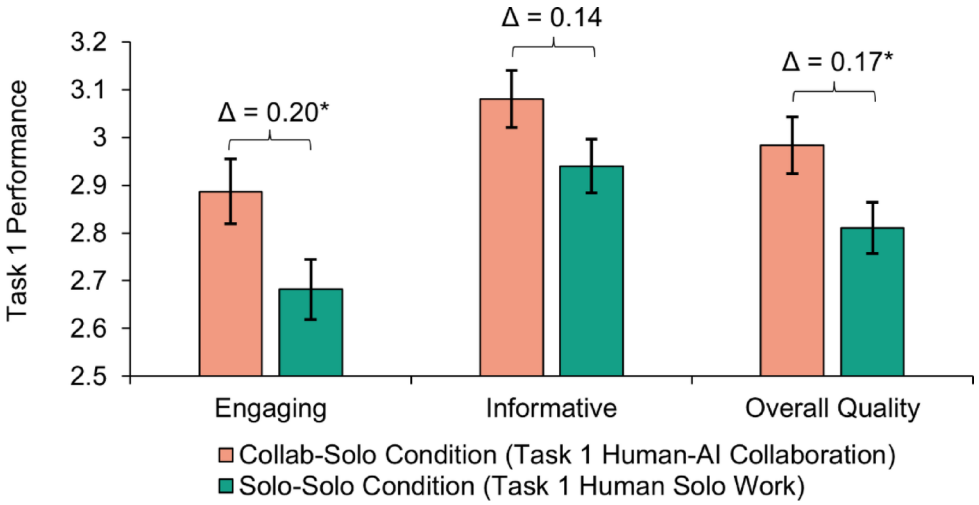


Fig. 3. Augmentation effect of collaboration with ChatGPT for Study 1, Task 1 performance. Independent raters assessed how engaging and informative the posts created by participants were. The overall quality score was computed as the means of these two ratings. In contrast to individuals who created Facebook posts on their own, those who collaborated with ChatGPT produced posts that were more engaging and of higher overall quality. Nevertheless, there was no significant difference between the two participant groups in terms of post informativeness. $N = 352$. $^{*}p < .05$. Error bars indicate ± 1 SEM.

Our focus was to discern whether there were differences in both the quantity (i.e., number of ideas) and quality (i.e., overall creativity) of outputs between the two groups. Data analysis revealed no significant difference in the quantity of ideas generated between the groups, with both demonstrating comparable performance in terms of idea quantity ($M_{\text{Collab–Solo}} = 7.46$, $SD_{\text{Collab–Solo}} = 3.47$ vs. $M_{\text{Solo–Solo}} = 7.39$, $SD_{\text{Solo–Solo}} = 4.03$; $t(350) = 0.17$, $p = .865$, 95% $CI = [-0.72, 0.86]$, $d = 0.02$). However, a significant difference emerged in the qualitative assessment. Participants who collaborated with GenAI in the initial task produced ideas of notably higher creativity in the subsequent task compared to those who worked independently throughout ($M_{\text{Collab–Solo}} = 3.34$, $SD_{\text{Collab–Solo}} = 0.55$ vs. $M_{\text{Solo–Solo}} = 3.17$, $SD_{\text{Solo–Solo}} = 0.61$; $t(350) = 2.73$, $p = .007$, 95% $CI = [0.05, 0.29]$, $d = 0.29$). This suggests mixed findings of ChatGPT's augmentation effect on the subsequent task performance. ChatGPT continued to augment the subsequent human-solo task (Task 2) performance on idea creativity (small effect size) but not on idea quantity (negligible effect size).

Deprivation effect of human-GenAI collaboration on psychological experiences (RQ2-4)

How does transitioning from human-GenAI collaboration to solo work change individuals' psychological experiences? To investigate this, we conducted mixed-design repeated measures ANOVAs to examine the interaction effects between task (Task 1 vs. Task 2) and condition (Collab-Solo vs. Solo-Solo) on participants' sense of control, intrinsic motivation, and feelings of boredom, respectively. Additionally, paired-samples *t*-tests were used to assess within-individual changes in psychological experiences from Task 1 to Task 2 (see Table 3; Fig. 5 for a summary of the results).

DV	Quantity of Ideas		Quality of Ideas	
Condition	Collab-Solo	Solo-Solo	Collab-Solo	Solo-Solo
Mean	7.46	7.39	3.34	3.17
SD	3.47	4.03	0.55	0.61
t (*df*)	0.17 (350)		**2.73 (350)**	
p	0.865		**0.007**	
95% *CI*	[-0.72, 0.86]		**[0.05, 0.29]**	
Cohen's *d*	0.02		**0.29**	

Table 2. Summary of main findings of Study 1, Task 2 performance across two conditions. DV = dependent variable, *SD* = standard deviation, *CI* = confidence interval, Cohen's *d* is the standardized effect size. Comparisons were conducted between Collab-Solo (transitioning from collaboration to solo work) and Solo-Solo (working solo in both tasks) conditions. *T*-tests were used to evaluate mean differences for each dependent variable. Significant results are bolded for emphasis.

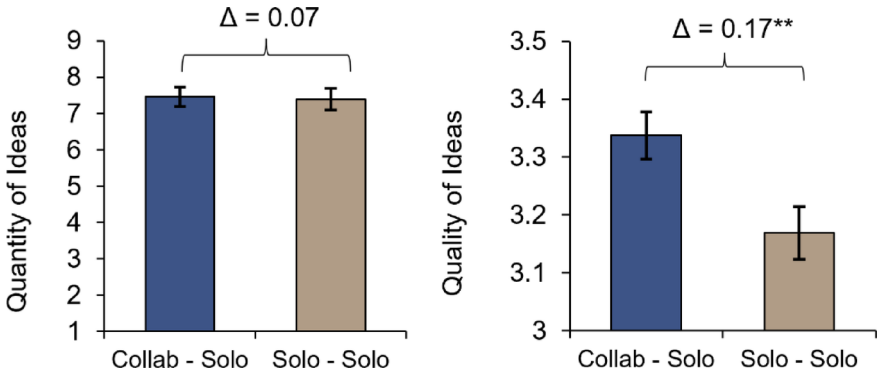


Fig. 4. Study 1, Task 2 performance across two conditions. Independent raters counted the number of ideas; and assessed the quality of these ideas on a five-point scale with endpoints 1 for "poor" and 5 for "excellent" in Task 2 (Alternative Uses Test; AUT). Participants who collaborated with GenAI in Task 1 generated ideas of higher quality in the subsequent human-solo work in Task 2, compared to participants who worked solo in both tasks. No significant difference was found between the two groups in terms of the number of ideas generated in Task 2. $N = 352$. $^{**}p < .01$. Error bars indicate ± 1 SEM.

DV	Sense of Control				Intrinsic Motivation				Feelings of Boredom			
Condition	Collab-Solo		Solo-Solo		Collab-Solo		Solo-Solo		Collab-Solo		Solo-Solo	
Task	Task 1	Task 2	Task 1	Task 2	Task 1	Task 2	Task 1	Task 2	Task 1	Task 2	Task 1	Task 2
Mean	5.57	5.78	5.67	5.52	5.08	4.39	4.86	4.62	3.43	3.91	3.42	3.64
SD	1.1	1.12	1.03	1.28	1.25	1.46	1.41	1.48	0.98	0.97	0.95	1.08
t (*df*)	**-2.20 (175)**		1.65 (175)		**5.98 (175)**		**2.35 (175)**		**-5.20 (175)**		**-2.80 (175)**	
p	**0.029**		0.101		**< 0.001**		**0.02**		**< 0.001**		**0.006**	
95% *CI*	**[-0.38, -0.02]**		[-0.03, 0.35]		**[0.46, 0.91]**		**[0.04, 0.44]**		**[-0.66, -0.30]**		**[-0.36, -0.06]**	
Cohen's *d*	**0.19**		-0.13		**-0.51**		**-0.17**		**0.49**		**0.22**	
F (df1, df2)	**7.36 (1, 350)**				**8.41 (1, 350)**				**4.98 (1, 350)**			
p	**0.007**				**0.004**				**0.026**			
partial η^2	**0.021**				**0.023**				**0.014**			
ω^2	**0.018**				**0.021**				**0.011**			

Table 3. Summary of main findings of Study 1 deprivation effect of collaboration with ChatGPT on psychological experiences. DV = dependent variable, *SD* = standard deviation, *CI* = confidence interval, Cohen's *d* is the standardized effect size, partial η^2 and ω^2 indicate effect size in repeated measures ANOVA. Paired samples *t*-tests were conducted to assess within-individual cross-task differences, while repeated measures ANOVA evaluated interaction effects. All significant results are bolded for emphasis.

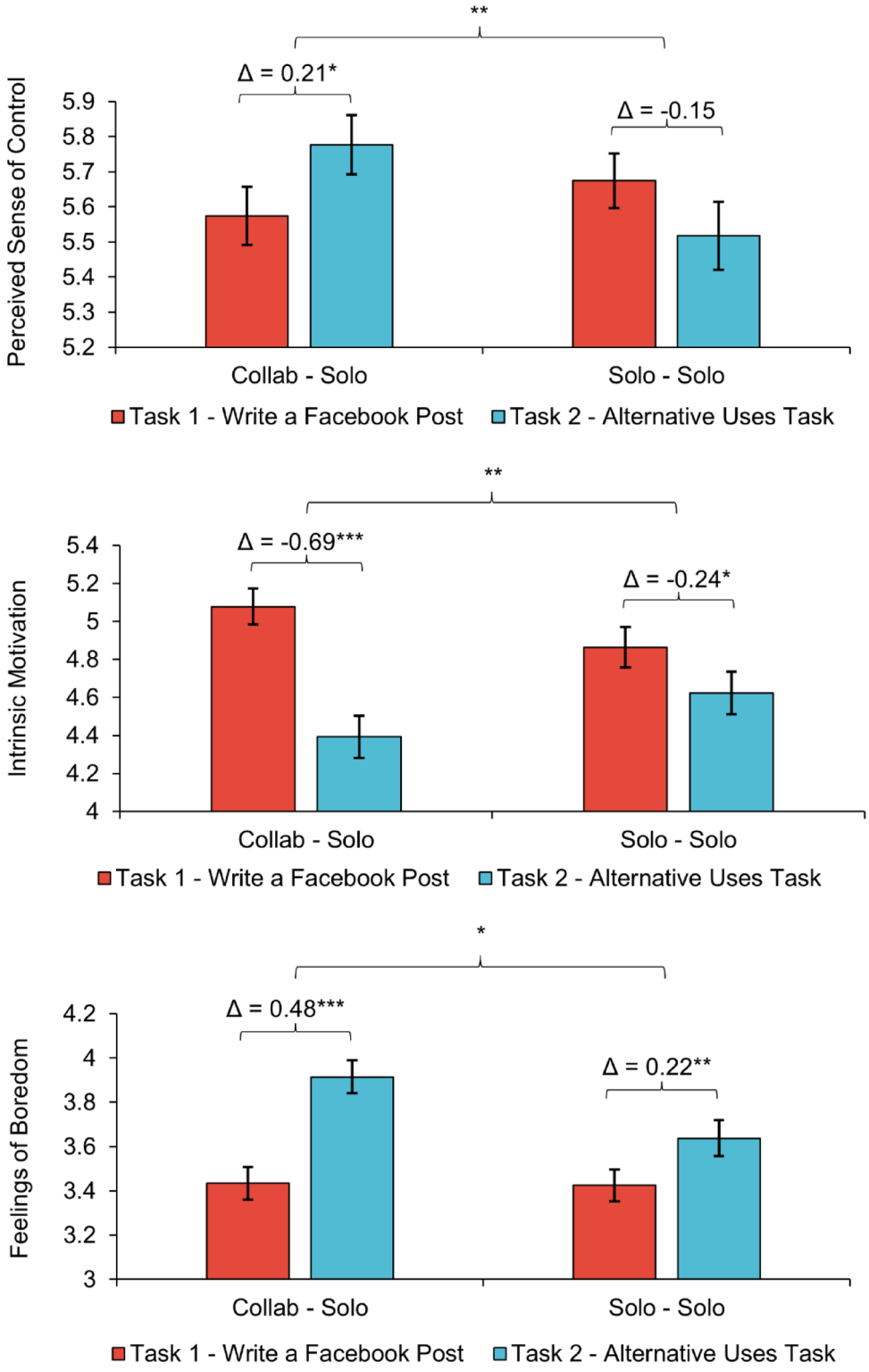


Fig. 5. Study 1, Deprivation effect of collaboration with ChatGPT on psychological experiences. Participants reported their sense of control, intrinsic motivation, and feelings of boredom on 7-point scales after Task 1 and Task 2, respectively. For participants who shifted from collaboration with ChatGPT to solo work, we observed a significant change in their sense of control (increased), intrinsic motivation (decreased), and feelings of boredom (increased). Such within-person changes were also significantly different from participants who consistently worked on two tasks alone. $N = 352$. $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$. Error bars indicate ± 1 SEM.

Increasing sense of control with task transition The analysis revealed a significant interaction effect between task and condition, $F(1, 350) = 7.36$, $p = .007$, partial $\eta^2 = 0.021$, $\omega^2 = 0.018$. For participants transitioning from human-GenAI collaboration to solo work, there was a significant increase (small effect size) in their sense of control from Task 1 ($M = 5.57$, $SD = 1.10$) to Task 2 ($M = 5.78$, $SD = 1.12$), $t(175) = -2.20$, $p = .029$, 95% $CI = [-0.38, -0.02]$, $d = 0.19$. This suggests that transitioning from collaborating with GenAI to solo work restores participants' sense of control. Their perceived sense of control may have been compromised during the initial collaboration with GenAI. On the other hand, participants who worked independently in both tasks showed a decrease in their sense of control from Task 1 ($M = 5.67$, $SD = 1.03$) to Task 2 ($M = 5.52$, $SD = 1.28$). Nevertheless, this decline was not statistically significant, $t(175) = 1.65$, $p = .101$, 95% $CI = [-0.03, 0.35]$, $d = -0.13$. Although this decline was not statistically significant, the negative effect size suggests a potential deprivation effect where consecutive human-solo tasks might diminish individuals' sense of control.

Diminishing intrinsic motivation with task transition The analysis revealed a significant interaction between task and condition, $F(1, 350) = 8.41$, $p = .004$, partial $\eta^2 = 0.023$, $\omega^2 = 0.021$. Participants who transitioned from human-GenAI collaboration in Task 1 to solo work in Task 2 exhibited a reduction in intrinsic motivation. The average score moves from 5.08 ($SD = 1.25$) in Task 1 to 4.39 ($SD = 1.46$) in Task 2, $t(175) = 5.98$, $p < .001$, 95% $CI = [0.46, 0.91]$, $d = -0.51$. As predicted, the removal of GenAI led to a notable decline (i.e., medium effect size) in human workers' intrinsic motivation in their subsequent solo tasks. In contrast, for those who performed both tasks independently, there was a modest decline (i.e., small effect size) in intrinsic motivation. Their scores dropped from 4.86 ($SD = 1.41$) in Task 1 to 4.62 ($SD = 1.48$) in Task 2, $t(175) = 2.35$, $p = .020$, 95% $CI = [0.04, 0.44]$, $d = -0.17$. This suggests a somewhat steady motivational experience for those consistently working without GenAI.

Surge in boredom with task transition A significant interaction between task (Task 1 vs. Task 2) and condition was observed, $F(1, 350) = 4.98$, $p = .026$, partial $\eta^2 = 0.014$, $\omega^2 = 0.011$. More specifically, participants who transitioned from collaboration in Task 1 to solo work in Task 2 showed an increase in boredom. Their scores went from 3.43 ($SD = 0.98$) in Task 1 to 3.91 ($SD = 0.97$) in Task 2, $t(175) = -5.20$, $p < .001$, 95% $CI = [-0.66, -0.30]$, $d = 0.49$. This medium effect size suggests that the removal of GenAI makes solo work significantly less engaging. Interestingly, participants who worked solo in both tasks also reported a slight increase in boredom. Their scores shifted from 3.42 ($SD = 0.95$) in Task 1 to 3.64 ($SD = 1.08$) in Task 2, $t(175) = -2.80$, $p = .006$, 95% $CI = [-0.36, -0.06]$, $d = 0.22$. The small effect size suggests that this mild increase in boredom could be attributed to the repetitive nature of performing two content-generation tasks consecutively.

Study 2

In Study 2 ($N = 793$), participants were first asked to draft a performance review report with or without ChatGPT's assistance (Collab-Solo vs. Solo-Solo, respectively) and filled in questions that assessed their psychological experience in Task 1, including sense of control, intrinsic motivation, and feeling of boredom. Then they were asked to brainstorm creative ideas for enhancing a product (i.e., an interactive whiteboard) on their own. After that, they again reported their psychological experience in Task 2.

Performance augmentation of ChatGPT (RQ1)

In this study, we utilized the Linguistic Inquiry and Word Count (LIWC)[41] software to evaluate performance review reports under two conditions: collaboration with ChatGPT and solo drafting. The analysis revealed that reports written in collaboration with ChatGPT were significantly longer, with a mean word count of 289.76 ($SD = 154.26$) compared to solo reports at 108.51 ($SD = 74.69$; $t(582.78) = 21.15$, $p < .001$, 95% $CI = [164.42, 198.09]$, $d = 1.50$). This large effect size highlights a substantial difference in productivity between the two conditions. Analytical content was also higher in collaborative reports ($M = 68.05$, $SD = 16.10$) versus solo ($M = 63.11$, $SD = 23.48$; $t(688.52) = 3.44$, $p = .001$, 95% $CI = [2.12, 7.75]$, $d = 0.25$). This small effect size suggests a modest increase in the analytical depth of collaboratively written texts. Prosocial scores were greater in the collaborative condition ($M = 3.85$, $SD = 1.76$) than in solo work ($M = 3.05$, $SD = 1.75$; $t(791) = 6.45$, $p < .001$, 95% $CI = [0.56, 1.05]$, $d = 0.46$). This medium effect size indicates a meaningful improvement in the prosocial tone of collaborative outputs. These findings demonstrate how AI-assisted collaboration can influence different aspects of text production in performance reviews (see Table 4; Fig. 6 for the result summary).

Human-GenAI collaboration and the subsequent human-alone task performance (RQ1)

We further assessed whether participants from the human-GenAI collaboration condition outperformed their counterparts in the subsequent task (see Table 5; Fig. 7 for the result summary). In this task, participants were asked to generate as many ideas as possible for improving a product, without any AI assistance. The analysis aims to test whether there were differences in both the quantity (i.e., number of ideas) and quality (i.e., novelty and usefulness) of outputs between the two groups. Data analysis revealed no significant difference in the quantity of ideas generated between the groups, with both demonstrating comparable performance in terms of idea quantity ($M_{\text{Collab–Solo}} = 5.26$, $SD_{\text{Collab–Solo}} = 3.70$ vs. $M_{\text{Solo–Solo}} = 5.14$, $SD_{\text{Solo–Solo}} = 3.25$; $t(791) = 0.48$, $p = .634$, 95% $CI = [-0.37, 0.60]$, $d = 0.03$), idea novelty ($M_{\text{Collab–Solo}} = 2.28$, $SD_{\text{Collab–Solo}} = 0.90$ vs. $M_{\text{Solo–Solo}} = 2.23$, $SD_{\text{Solo–Solo}} = 0.83$; $t(791) = 0.78$, $p = .436$, 95% $CI = [-0.07, 0.17]$, $d = 0.06$), and idea usefulness ($M_{\text{Collab–Solo}} = 2.49$, $SD_{\text{Collab–Solo}} = 0.84$ vs. $M_{\text{Solo–Solo}} = 2.47$, $SD_{\text{Solo–Solo}} = 0.77$; $t(791) = 0.26$, $p = .794$, 95% $CI = [-0.10, 0.13]$, $d = 0.02$). These negligible effect sizes indicate that ChatGPT did not extend the augmentation effect on participants' immediate performance to the subsequent human-solo task.

DV	Word Count		Analytic Content		Prosocial Orientation	
Condition	Collab-Solo	Solo-Solo	Collab-Solo	Solo-Solo	Collab-Solo	Solo-Solo
Mean	289.76	108.51	68.05	63.11	3.85	3.05
SD	154.26	74.69	16.10	23.48	1.76	1.75
t (*df*)	**21.15 (582.78)**		**3.44 (688.52)**		**6.45 (791)**	
p	**< 0.001**		**0.001**		**< 0.001**	
95% *CI*	**[164.42, 198.09]**		**[2.12, 7.75]**		**[0.56, 1.05]**	
Cohen's *d*	**1.50**		**0.25**		**0.46**	

Table 4. Summary of main findings of augmentation effect of collaboration with ChatGPT for Study 2, Task 1 performance. DV = dependent variable, *SD* = standard deviation, *CI* = confidence interval, Cohen's *d* is the standardized effect size. Comparisons were conducted between Collab-Solo (transitioning from collaboration to solo work) and Solo-Solo (working solo in both tasks) conditions. *T*-tests were used to evaluate mean differences for each dependent variable. Significant results are bolded for emphasis.

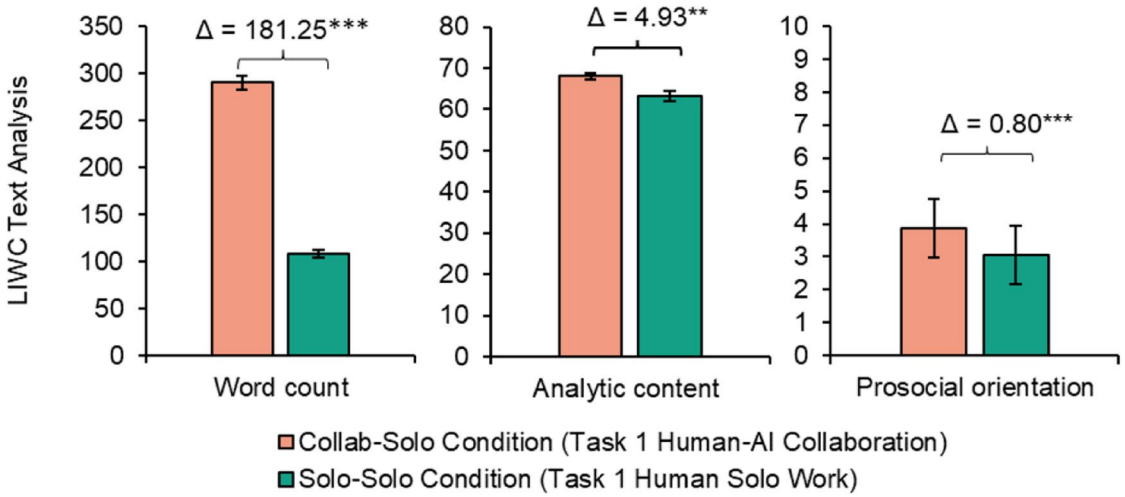


Fig. 6. Augmentation effect of collaboration with ChatGPT for Study 2, Task 1 performance. We analyzed the quality of performance review reports across three dimensions—word count, analytical content, and prosocial orientation—using the LIWC tool. Results suggest collaborating with ChatGPT augments participants' Task 1 performance. Participants collaborating with ChatGPT produced outputs of higher quality across all three dimensions: They had greater word counts, exhibited more analytical content, and demonstrated enhanced prosocial orientation compared to those working without ChatGPT's assistance. $N = 793$. $^{**}p < .01$; $^{***}p < .001$. Error bars indicate ± 1 SEM.

DV	Idea Quantity		Idea Novelty		Idea Usefulness	
Condition	Collab-Solo	Solo-Solo	Collab-Solo	Solo-Solo	Collab-Solo	Solo-Solo
Mean	5.26	5.14	2.28	2.23	2.49	2.47
SD	3.70	3.25	0.90	0.83	0.84	0.77
t (*df*)	0.48 (791)		0.78 (791)		0.26 (791)	
p	0.634		0.436		0.794	
95% *CI*	[-0.37, 0.60]		[-0.07, 0.17]		[-0.10, 0.13]	
Cohen's *d*	0.03		0.06		0.02	

Table 5. Summary of main findings of Study 2, Task 2 performance across two conditions. *DV* = dependent variable, *SD* = standard deviation, *CI* = confidence interval, Cohen's *d* is the standardized effect size. Comparisons were conducted between Collab-Solo (transitioning from collaboration to solo work) and Solo-Solo (working solo in both tasks) conditions. *T*-tests were used to evaluate mean differences for each dependent variable. Significant results are bolded for emphasis.

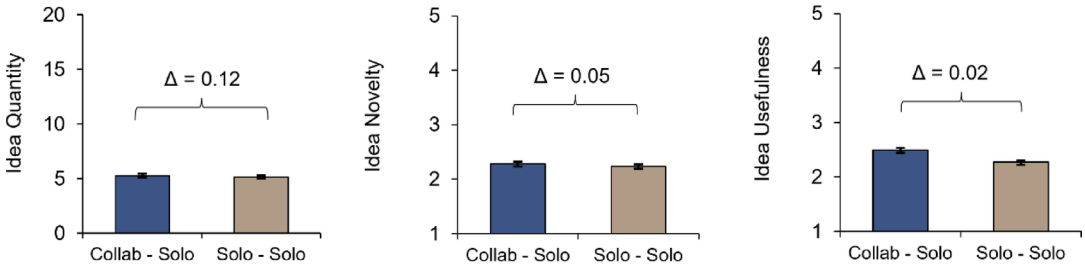


Fig. 7. Study 2, Task 2 performance across two conditions. Independent raters counted the number of ideas and assessed their novelty and usefulness on a five-point scale (1 = "poor", 5 = "excellent") in Task 2, where participants brainstormed ideas for improving a product. Analysis of data showed no significant differences between the two conditions in terms of idea quantity, novelty, or usefulness, suggesting that collaboration with ChatGPT did not significantly augment performance in the subsequent solo task. $N = 793$. $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$. Error bars indicate ± 1 SEM.

DV	Sense of Control				Intrinsic Motivation				Feelings of Boredom			
Condition	Collab-Solo		Solo-Solo		Collab-Solo		Solo-Solo		Collab-Solo		Solo-Solo	
Task	Task 1	Task 2	Task 1	Task 2	Task 1	Task 2	Task 1	Task 2	Task 1	Task 2	Task 1	Task 2
Mean	5.55	5.72	5.90	5.36	4.85	4.38	4.67	4.34	2.72	3.36	2.75	3.18
SD	1.16	1.13	0.98	1.35	1.36	1.56	1.42	1.53	1.26	1.57	1.30	1.46
t (*df*)	**-2.34 (401)**		**7.53 (390)**		**5.75 (401)**		**4.06 (390)**		**-7.96 (401)**		**-5.65 (390)**	
p	**0.020**		**< 0.001**		**< 0.001**		**< 0.001**		**< 0.001**		**< 0.001**	
95% *CI*	**[-0.31, -0.03]**		**[0.40, 0.68]**		**[0.31, 0.63]**		**[0.17, 0.49]**		**[-0.80, -0.48]**		**[-0.58, -0.28]**	
Cohen's *d*	**0.15**		**-0.46**		**-0.32**		**-0.22**		**0.45**		**0.31**	
F (df1, df2)	**48.46 (1, 791)**				1.47 (1, 791)				3.52 (1, 791)			
p	**< 0.001**				0.225				0.061			
partial η^2	**0.058**				0.002				0.004			
ω^2	**0.057**				0.001				0.003			

Table 6. Summary of main findings of Study 2 deprivation effect of collaboration with ChatGPT on psychological experiences. DV = dependent variable, *SD* = standard deviation, *CI* = confidence interval, Cohen's *d* is the standardized effect size, partial η^2 and ω^2 indicate effect size in repeated measures ANOVA. Paired samples *t*-tests were conducted to assess within-individual cross-task differences, while repeated measures ANOVA evaluated interaction effects. All significant results are bolded for emphasis.

Deprivation effect of human-GenAI collaboration on psychological experiences (RQ2-4)
Similar to Study 1, we examined how transitioning from human-GenAI collaboration to solo work (compared to transitioning from solo work to solo work) changed individuals' psychological experiences. Participants reported their sense of control, intrinsic motivation, and feelings of boredom after Task 1 and Task 2, respectively. Following similar analytical strategies as in Study 1, we conducted mixed-design repeated measures ANOVAs to examine the interaction effects between task (Task 1 vs. Task 2) and condition (Collab-Solo vs. Solo-Solo) on participants' sense of control, intrinsic motivation, and feelings of boredom, respectively. Additionally, paired-samples *t*-tests were used to assess within-individual changes in psychological experiences from Task 1 to Task 2 (see Table 6; Fig. 8 for the result summary).

Increasing sense of control with task transition Results revealed a significant interaction effect between task and condition, $F(1, 791) = 48.46$, $p < .001$, partial $\eta^2 = 0.058$, $\omega^2 = 0.057$. For participants transitioning from human-GenAI collaboration to solo work, there was a significant increase in their sense of control from Task 1 ($M = 5.55$, $SD = 1.16$) to Task 2 ($M = 5.72$, $SD = 1.13$), $t(401) = -2.34$, $p = .020$, 95% $CI = [-0.31, -0.03]$, $d = 0.15$. This small effect size suggests a modest regain in perceived control when transitioning away from collaboration with GenAI. On the other hand, participants who worked alone in both tasks showed a significant decrease in their sense of control from Task 1 ($M = 5.90$, $SD = 0.98$) to Task 2 ($M = 5.36$, $SD = 1.35$), $t(390) = 7.53$, $p < .001$, 95% $CI = [0.40, 0.68]$, $d = -0.46$. This medium negative effect size indicates a notable decline in their sense of control across the two tasks. This pattern implies that the solo work following collaboration with GenAI may have contributed to a notable regain in perceived control, while those consistently working alone may have faced a notable decline in their sense of control across the two tasks.

Diminishing intrinsic motivation across tasks Results indicated that the interaction between task and condition was not significant ($F(1, 791) = 1.47$, $p = .225$, partial $\eta^2 = 0.002$, $\omega^2 = 0.001$). Participants who transitioned

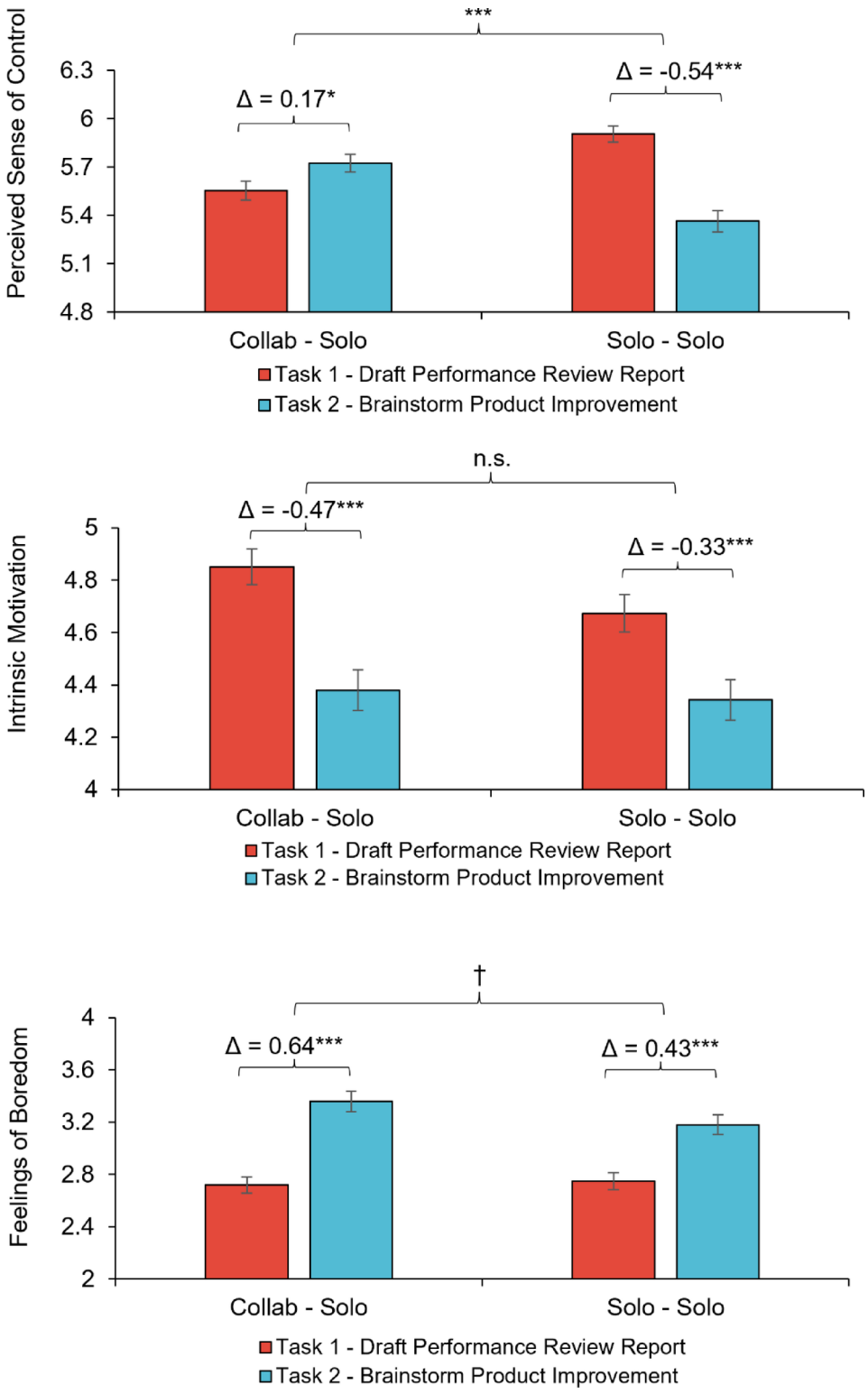


Fig. 8. Study 2, Deprivation effect of collaboration with ChatGPT on psychological experiences. Participants transitioning from collaboration with ChatGPT to solo work reported a significant increase in their sense of control, while those who worked independently on both tasks reported a significant decrease. Additionally, participants in both conditions experienced a notable decrease in intrinsic motivation and an increase in feelings of boredom. However, the differences in these changes between the two conditions were non-significant for intrinsic motivation and only marginally significant for feelings of boredom. $N = 793$. $^{\dagger}p < .10$, $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$. Error bars indicate ± 1 SEM.

from human-GenAI collaboration in Task 1 to solo work in Task 2 exhibited a marked reduction in intrinsic motivation from Task 1 ($M = 4.85$, $SD = 1.36$) to Task 2 ($M = 4.38$, $SD = 1.56$), $t(401) = 5.75$, $p < .001$, 95% CI = [0.31, 0.63], d = -0.32. This medium negative effect size highlights a significant decline in motivation following the removal of GenAI assistance. Similarly, there was also a significant decline in intrinsic motivation for those who engaged in both tasks alone, dropping from 4.67 ($SD = 1.42$) in Task 1 to 4.34 ($SD = 1.53$) in Task 2, $t(390) = 4.06$, $p < .001$, 95% CI = [0.17, 0.49], d = -0.22. The small negative effect size suggests a less pronounced but still notable decline in motivation for this group. Although the interaction between task types and conditions did not reach statistical significance, the clear decline in intrinsic motivation across all participants is notable. This decline was more pronounced in individuals transitioning from human-GenAI collaboration to solo tasks, suggesting that while the GenAI's initial assistance might boost task performance, it does not necessarily sustain motivational levels in subsequent independent work.

Increasing feelings of boredom across tasks A marginally significant interaction between task (Task 1 vs. Task 2) and condition was observed, $F(1, 791) = 3.52$, $p = .061$, partial $\eta^2 = 0.004$, $\omega^2 = 0.003$. Participants who transitioned from collaboration in Task 1 to solo work in Task 2 showed an increase in feelings of boredom. Their scores went from 2.72 ($SD = 1.26$) in Task 1 to 3.36 ($SD = 1.57$) in Task 2, $t(401)$ = -7.96, $p < .001$, 95% CI = [-0.80, -0.48], $d = 0.45$. This medium effect size suggests a notable increase in boredom following the transition to solo work. Participants who worked solo in both tasks also reported an increase in feelings of boredom. Their scores increased from 2.75 ($SD = 1.30$) in Task 1 to 3.18 ($SD = 1.46$) in Task 2, $t(390)$ = -5.65, $p < .001$, 95% CI = [-0.58, -0.28], $d = 0.31$. This small-to-medium effect size indicates a more modest increase in boredom for participants consistently working alone. Although the interaction effect was not statistically significant at the conventional threshold, the observed patterns suggest a possible influence of task type and collaboration conditions on boredom levels. Specifically, both groups experienced increased boredom over time, with a more pronounced increase among participants who initially collaborated with GenAI. This suggested that human-GenAI collaboration has a potentially stronger effect on the increase in feelings of boredom when preceded to human solo tasks.

Study 3

In Study 3 ($N = 793$), participants were instructed to take the role of a team manager and write an email introducing a new colleague to the whole team with or without the assistance from ChatGPT (Collab-Solo vs. Solo-Solo, respectively) and then were asked to report their psychological experience in Task 1, including sense of control, intrinsic motivation, and feeling of boredom. Next, participants proceeded with Task 2, which asked them to generate creative marketing ideas for a specified cleaning product by themselves. After that, participants again reported their psychological experiences in Task 2.

Performance augmentation of ChatGPT (RQ1)

In Study 3, we also used the Linguistic Inquiry and Word Count (LIWC) software to analyze emails welcoming new colleagues under two conditions: collaboration with ChatGPT and solo drafting. The analysis showed that emails written with the aid of ChatGPT were significantly longer, with a mean word count of 201.66 ($SD = 82.30$) compared to independently drafted emails at 111.22 ($SD = 59.58$; $t(736.70) = 17.78$, $p < .001$, 95% CI = [80.45, 100.42], $d = 1.26$). This large effect size indicates a substantial increase in email length when collaborating with GenAI. Emails drafted collaboratively with ChatGPT also scored higher in affiliation content (M_{Collab} = 11.83, SD_{Collab} = 4.07) compared to those drafted independently (M_{Solo} = 9.89, SD_{Solo} = 3.77; $t(791) = 6.96$, $p < .001$, 95% CI = [1.39, 2.49], $d = 0.49$). This medium effect size suggests meaningful improvements in the affiliative tone of emails when GenAI is used. Finally, emails written collaboratively were more socially expressive (M_{Collab} = 22.54, SD_{Solo} = 4.82 vs. M_{Solo} = 21.42, SD_{Solo} = 4.82; $t(791) = 3.27$, $p = .001$, 95% CI = [0.45, 1.79], $d = 0.23$). This small effect size indicates modest enhancements in social expressiveness. These findings highlight how GenAI-collaboration enhances both the engagement and social warmth of email communications (see Table 7; Fig. 9 for the result summary).

DV	Word Count		Affiliation Content		Social Orientation	
Condition	Collab-Solo	Solo-Solo	Collab-Solo	Solo-Solo	Collab-Solo	Solo-Solo
Mean	201.66	111.22	11.83	9.89	22.54	21.42
SD	82.30	59.58	4.07	3.77	4.82	4.82
t (*df*)	**17.78 (736.70)**		**6.96 (791)**		**3.27 (791)**	
p	**< 0.001**		**< 0.001**		**0.001**	
95% *CI*	**[80.45, 100.42]**		**[1.39, 2.49]**		**[0.45, 1.79]**	
Cohen's *d*	**1.26**		**0.49**		**0.23**	

Table 7. Summary of main findings of augmentation effect of collaboration with ChatGPT for Study 3, Task 1 performance. DV = dependent variable, *SD* = standard deviation, *CI* = confidence interval, Cohen's *d* is the standardized effect size. Comparisons were conducted between Collab-Solo (transitioning from collaboration to solo work) and Solo-Solo (working solo in both tasks) conditions. *T*-tests were used to evaluate mean differences for each dependent variable. Significant results are bolded for emphasis.

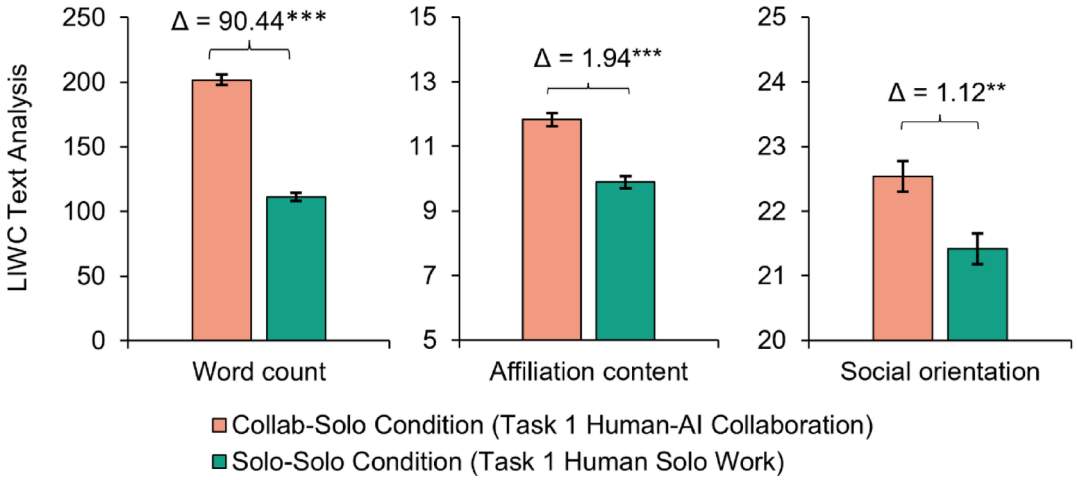


Fig. 9. Augmentation effect of collaboration with ChatGPT for Study 3, Task 1 performance. We analyzed the quality of welcoming emails using LIWC across three dimensions: Word count, affiliation content, and social orientation. Results indicated that collaboration with ChatGPT enhanced Task 1 performance, with participants who used ChatGPT producing outputs that were not only longer but also demonstrated greater affiliation content and higher social orientation compared to those who worked independently. $N = 793$. $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$. Error bars indicate ± 1 SEM.

DV	Idea Quantity		Idea Novelty		Idea Usefulness	
Condition	Collab-Solo	Solo-Solo	Collab-Solo	Solo-Solo	Collab-Solo	Solo-Solo
Mean	4.72	4.91	2.48	2.56	2.52	2.59
SD	2.77	2.92	0.87	0.86	0.75	0.74
t (*df*)	-0.95 (791)		-1.37 (791)		-1.48 (791)	
p	0.344		0.170		0.140	
95% *CI*	[-0.56, 0.21]		[-0.20, 0.04]		[-0.18, 0.03]	
Cohen's *d*	-0.07		-0.09		-0.09	

Table 8. Summary of main findings of Study 3, Task 2 performance across two conditions. DV = dependent variable, *SD* = standard deviation, *CI* = confidence interval, Cohen's *d* is the standardized effect size. Comparisons were conducted between Collab-Solo (transitioning from collaboration to solo work) and Solo-Solo (working solo in both tasks) conditions. *T*-tests were used to evaluate mean differences for each dependent variable. Significant results are bolded for emphasis.

Human-GenAI collaboration and the subsequent human-alone task performance (RQ1)

We further assessed whether participants from the human-GenAI collaboration condition outperformed their counterparts in the subsequent task. In this task, all participants were tasked with generating as many product promotion ideas as possible without any AI assistance. The analysis aims to test if there were differences in both the quantity (i.e., number of ideas) and quality (i.e., novelty and usefulness) of outputs between the two groups. Data analysis revealed no significant difference in the quantity of ideas generated between the groups, with both demonstrating comparable performance in terms of idea quantity ($M_{\text{Collab–Solo}} = 4.72$, $SD_{\text{Collab–Solo}} = 2.77$ vs. $M_{\text{Solo–Solo}} = 4.91$, $SD_{\text{Solo–Solo}} = 2.92$; $t(791) = -0.95$, $p = .344$, 95% $CI = [-0.56, 0.21]$, $d = -0.07$), idea novelty ($M_{\text{Collab–Solo}} = 2.48$, $SD_{\text{Collab–Solo}} = 0.87$ vs. $M_{\text{Solo–Solo}} = 2.56$, $SD_{\text{Solo–Solo}} = 0.86$; $t(791) = -1.37$, $p = .170$, 95% $CI = [-0.20, 0.04]$, $d = -0.09$), and idea usefulness ($M_{\text{Collab–Solo}} = 2.52$, $SD_{\text{Collab–Solo}} = 0.75$ vs. $M_{\text{Solo–Solo}} = 2.59$, $SD_{\text{Solo–Solo}} = 0.74$; $t(791) = -1.48$, $p = .140$, 95% $CI = [-0.18, 0.03]$, $d = -0.09$). These negligible effect sizes indicate that the augmentation effect of ChatGPT on participants' immediate performance did not carry over to the subsequent human-solo task (see Table 8; Fig. 10 for the results summary).

Deprivation effect of human-GenAI collaboration on psychological experiences (RQ2-4)

Again, we examined how transitioning from human-GenAI collaboration to solo work (compared to transitioning from solo work to solo work) changed individuals' psychological experiences. Participants reported their sense of control, intrinsic motivation, and feelings of boredom after Task 1 and Task 2, respectively. Following similar analytical strategies as in Study 1, we conducted mixed-design repeated measures ANOVAs to examine the interaction effects between task (Task 1 vs. Task 2) and condition (Collab-Solo vs. Solo-Solo) on participants' sense of control, intrinsic motivation, and feelings of boredom, respectively. Additionally, paired-samples *t*-tests

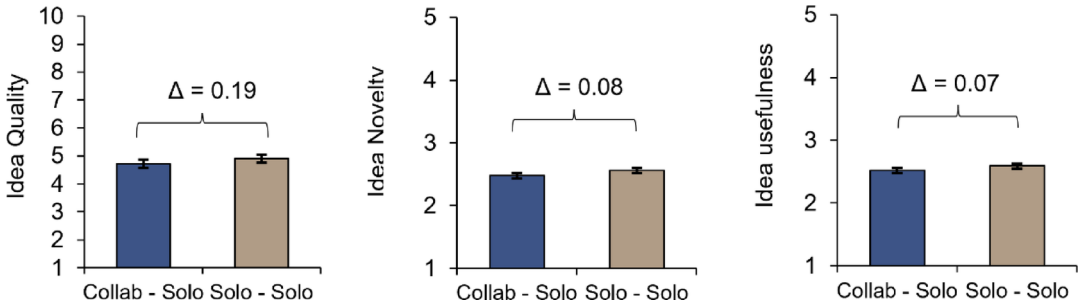


Fig. 10. Study 3, Task 2 performance across two conditions. Independent raters counted the number of ideas; and evaluated their novelty and usefulness on a five-point scale (1 = "poor", 5 = "excellent") in Task 2, where participants brainstormed promotional strategies for a product. The analysis revealed no significant differences between the two conditions in terms of the quantity, novelty, or usefulness of the ideas, indicating that collaboration with ChatGPT did not significantly enhance performance in the subsequent human-solo task. $N = 793$. $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$. Error bars indicate ± 1 SEM.

DV	Sense of Control				Intrinsic Motivation				Feelings of Boredom			
Condition	Collab-Solo		Solo-Solo		Collab-Solo		Solo-Solo		Collab-Solo		Solo-Solo	
Task	Task 1	Task 2	Task 1	Task 2	Task 1	Task 2	Task 1	Task 2	Task 1	Task 2	Task 1	Task 2
Mean	5.76	5.88	5.96	5.51	5.10	4.48	4.95	4.58	2.52	3.21	2.63	3.10
SD	1.01	1.06	1.02	1.20	1.29	1.52	1.24	1.56	1.20	1.49	1.17	1.52
t (*df*)	-1.82 (404)		**7.01 (387)**		**7.65 (404)**		**4.84 (387)**		**-8.50 (404)**		**-6.48 (387)**	
p	0.070		**< 0.001**		**< 0.001**		**< 0.001**		**< 0.001**		**< 0.001**	
95% *CI*	[-0.24, 0.01]		**[0.33, 0.58]**		**[0.47, 0.79]**		**[0.22, 0.53]**		**[-0.85, -0.53]**		**[-0.61, -0.33]**	
Cohen's *d*	0.12		**-0.40**		**-0.44**		**-0.26**		**0.51**		**0.35**	
F (df1, df2)	**39.09 (1, 791)**				**5.05 (1, 791)**				**4.08 (1, 791)**			
p	**< 0.001**				**0.025**				**0.044**			
partial η^2	**0.047**				**0.006**				**0.005**			
ω^2	**0.046**				**0.005**				**0.004**			

Table 9. Summary of main findings of Study 3 deprivation effect of collaboration with ChatGPT on psychological experiences. *DV* = dependent variable, *SD* = standard deviation, *CI* = confidence interval, Cohen's *d* is the standardized effect size, partial η^2 and ω^2 indicate effect size in repeated measures ANOVA. Paired samples *t*-tests were conducted to assess within-individual cross-task differences, while repeated measures ANOVA evaluated interaction effects. All significant results are bolded for emphasis.

were used to assess within-individual changes in psychological experiences from Task 1 to Task 2 (see Table 9; Fig. 11 for the result summary).

Increasing sense of control with task transition Results revealed a significant interaction effect between task and condition, $F(1, 791) = 39.09$, $p < .001$, partial $\eta^2 = 0.047$, $\omega^2 = 0.046$. For participants transitioning from human-GenAI collaboration to solo work, there was a marginally significant increase in their sense of control from Task 1 ($M = 5.76$, $SD = 1.01$) to Task 2 ($M = 5.88$, $SD = 1.06$), $t(404) = -1.82$, $p = .070$, 95% $CI = [-0.24, 0.01]$, $d = 0.12$. This small positive effect size suggests a slight boost in perceived control during the transition to solo work. On the other hand, participants who worked alone in both tasks showed a significant decrease in their sense of control from Task 1 ($M = 5.96$, $SD = 1.02$) to Task 2 ($M = 5.51$, $SD = 1.20$), $t(387) = 7.01$, $p < .001$, 95% $CI = [0.33, 0.58]$, $d = -0.40$. This medium negative effect size reflects a notable decline in perceived control for participants consistently working alone. This pattern implies that the solo work following collaboration with GenAI may have contributed to a slight boost in perceived control, while those consistently working alone may have faced a notable decline in their sense of control across the two tasks.

Diminishing intrinsic motivation with task transition The analysis revealed a significant interaction between task and condition, $F(1, 791) = 5.05$, $p = .025$, partial $\eta^2 = 0.006$, $\omega^2 = 0.005$. Participants transitioning from human-GenAI collaboration in Task 1 to human-solo work in Task 2 experienced a significant drop in intrinsic motivation, from an average score of 5.10 ($SD = 1.29$) in Task 1 to 4.48 ($SD = 1.52$) in Task 2, $t(404) = 7.65$, $p < .001$, 95% $CI = [0.47, 0.79]$, $d = -0.44$. This medium negative effect size indicates a notable decline in motivation following the removal of GenAI assistance. Participants working independently on both tasks also showed a decrease in motivation, with scores falling from 4.95 ($SD = 1.24$) in Task 1 to 4.58 ($SD = 1.56$) in Task 2, $t(387) = 4.84$, $p < .001$, 95% $CI = [0.22, 0.53]$, $d = -0.26$. This small-to-medium effect size suggests a less pro-

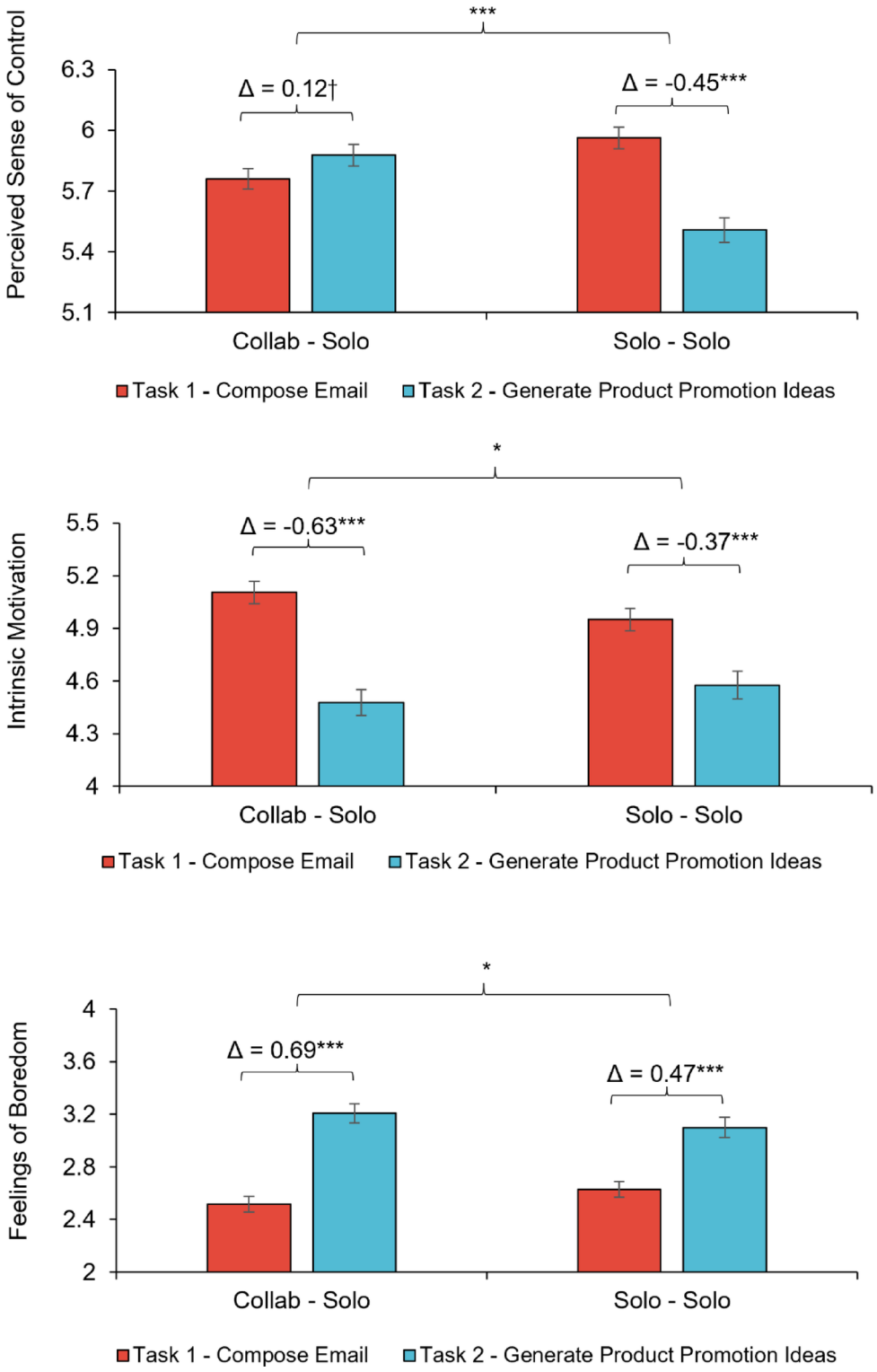


Fig. 11. Study 3, Deprivation effect of collaboration with ChatGPT on psychological experiences. Participants reported their sense of control, intrinsic motivation, and feelings of boredom on 7-point scales after Task 1 and Task 2, respectively. Those who transitioned from collaboration with ChatGPT to human-solo work showed a marginally significant increase in their sense of control. In contrast, participants who worked solo in both tasks experienced a significant decrease in their sense of control. Additionally, those shifting from collaboration to human-solo work reported significantly decreased intrinsic motivation and increased feelings of boredom, with these changes being significantly different from those who consistently worked solo on both tasks. $N = 352$. $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$. Error bars indicate ± 1 SEM.

nounced decline compared to the collaboration-to-solo transition group. These findings highlight how removing GenAI assistance significantly reduces intrinsic motivation in subsequent solo tasks, with a larger impact observed among participants transitioning from human-GenAI collaboration.

Increase in feelings of boredom with task transition A significant interaction between task (Task 1 vs. Task 2) and condition was observed, $F(1, 791) = 4.08$, $p = .044$, partial $\eta^2 = 0.005$, $\omega^2 = 0.004$. Participants who transitioned from collaboration in Task 1 to human-solo work in Task 2 experienced an increase in boredom, with scores rising from 2.52 ($SD = 1.20$) in Task 1 to 3.21 ($SD = 1.49$) in Task 2, $t(404) = -8.50$, $p < .001$, 95% $CI = [-0.85, -0.53]$, $d = 0.51$. This medium effect size reflects a notable increase in boredom following the transition to solo work. Participants who worked solo on both tasks also reported increased boredom, with scores increasing from 2.63 ($SD = 1.17$) in Task 1 to 3.10 ($SD = 1.52$) in Task 2, $t(387) = -6.48$, $p < .001$, 95% $CI = [-0.61, -0.33]$, $d = 0.35$. This small-to-medium effect size indicates a milder increase in boredom for participants consistently working alone. These findings suggest that both conditions experienced increased boredom over time, with the effect being more pronounced among participants transitioning from human-GenAI collaboration to solo tasks.

Study 4

Studies 1–3 examined the effects of two task transition conditions, Collab-Solo and Solo-Solo, on participants' psychological experiences and task performance. However, these studies did not explore two additional possible conditions: Maintaining collaboration with GenAI for both tasks and transitioning from solo work to collaborating with GenAI. Including these conditions would provide deeper insights into whether the observed effects stem from differences between working independently versus with GenAI or from the transition between GenAI-assisted and solo tasks. To address this gap, we conducted Study 4 ($N = 1{,}624$) using a 2×2 mixed factorial design. The study included one between-subjects factor (Collaboration with ChatGPT vs. Solo) and one within-subjects factor (Task 1 vs. Task 2), resulting in four experimental conditions: Collab-Collab ($N = 456$), Collab-Solo ($N = 345$), Solo-Solo ($N = 428$), and Solo-Collab ($N = 395$). Participants were randomly assigned to one of these four conditions.

As in Studies 1–3, participants in Study 4 performed two tasks consecutively and reported their psychological experiences—sense of control, intrinsic motivation, and feelings of boredom—after each task. However, unlike the earlier studies, Study 4 was designed to eliminate the influence of task type and task order by using two similar text-generation tasks and counterbalancing their sequence. Specifically, we used the "Composing an Email Task" from Study 3 and the "Composing a Facebook Post Task" from Study 1, presenting these tasks in random order to participants (see Fig. 2 for an overview of the experimental design). To ensure a comprehensive investigation, we analyzed data from all four conditions (Collab-Collab, Collab-Solo, Solo-Solo, and Solo-Collab) to examine the performance augmentation effect, the spillover of performance augmentation, and the deprivation effects on psychological experiences. By systematically comparing these conditions, Study 4 provides a more integrative examination of how transitioning between solo and AI-assisted work influences both task performance and psychological outcomes.

Performance augmentation of ChatGPT (RQ1)

We used the Linguistic Inquiry and Word Count (LIWC) tool to code participants' Task 1 outputs across three key dimensions relevant to text-generated tasks: text length (indicated by word count), analytical content, and positive tone. We compared the performance scores between participants who worked solo (including both the Solo-Solo and Solo-Collab conditions) and those who collaborated with GenAI (including the Collab-Solo and Collab-Collab conditions).

Results suggested that participants who worked with GenAI in Task 1 produced outputs that were longer ($M = 157.30$, $SD = 67.34$), more analytical ($M = 77.07$, $SD = 15.85$), and more positive ($M = 8.02$, $SD = 3.02$), whereas participants who worked solo in Task 1 generated outputs that were shorter ($M = 92.58$, $SD = 49.00$), less analytical ($M = 69.86$, $SD = 21.77$), and less positive ($M = 7.40$, $SD = 3.87$). These differences were significant for word count ($t(1458.77) = 22.10$, $p < .001$, 95% $CI = [58.97, 70.46]$, $d = 1.10$), analytical content ($t(1503.29) = 7.64$, $p < .001$, 95% $CI = [5.36, 9.06]$, $d = 0.38$), and positive tone ($t(1580.71) = 3.53$, $p < .001$, 95% $CI = [0.28, 0.97]$, $d = 0.17$). Taken together, these findings support the performance augmentation effect of ChatGPT, suggesting that collaboration with GenAI enhances task performance compared to working independently (see Table 10; Fig. 12 for a summary of results).

Human-GenAI collaboration and the subsequent human-solo task performance (RQ1)

We further assessed whether participants in the Collab-Solo condition outperformed their counterparts in the Solo-Solo condition during Task 2 across three key performance dimensions: text length, analytical content, and positive tone.

The analysis revealed a significant difference in Task 2 performance in terms of text length, with participants in the Solo-Solo condition producing slightly longer texts ($M_{\text{Solo–Solo}} = 88.47$, $SD_{\text{Solo–Solo}} = 50.02$) than those transitioning from collaboration ($M_{\text{Collab–Solo}} = 81.77$, $SD_{\text{Collab–Solo}} = 42.55$; $t(771) = -1.98$, $p = .048$, 95% $CI = [-13.35, -0.05]$, $d = 0.14$). This suggests that the performance augmentation observed during collaboration with GenAI did not carry over to subsequent solo tasks, as participants transitioning from collaboration performed slightly worse than those who worked solo throughout. In addition, no significant differences were observed for analytical content ($M_{\text{Collab–Solo}} = 74.32$, $SD_{\text{Collab–Solo}} = 21.27$ vs. $M_{\text{Solo–Solo}} = 72.21$, $SD_{\text{Solo–Solo}} = 21.68$; $t(771) = 1.35$, $p = .177$, 95% $CI = [-0.95, 5.16]$, $d = 0.10$), or positive tone ($M_{Collab–solo} = 7.10$, $SD_{Collab–solo} = 3.73$; $M_{Solo–solo} = 7.53$, $SD_{Solo–solo} = 4.18$; $t(762.94) = -1.53$, $p = .128$, 95% $CI = [-1.00, 0.12]$, $d = 0.11$). These negligible effect sizes suggest similar levels of analytical content and positive tone across the two groups. Overall, these findings indicate that

DV	Word Count				Analytical Content				Positive Tone			
Condition	Solo-Solo	Solo-Collab	Collab-Solo	Collab-Collab	Solo-Solo	Solo-Collab	Collab-Solo	Collab-Collab	Solo-Solo	Solo-Collab	Collab-Solo	Collab-Collab
Mean	91.43	93.83	150.16	162.7	69.97	69.75	76.35	77.62	7.43	7.37	8.08	7.79
SD	45.91	52.18	68.72	65.84	22.35	21.15	17.8	14.2	3.88	3.86	3.37	3.06
Mean	92.58		157.30		69.86		77.07		7.40		8.02	
SD	49.00		67.34		21.77		15.85		3.87		3.20	
t (df)	**22.10 (1459.77)**				**7.64 (1503.29)**				**3.53 (1580.71)**			
p	**<0.001**				**<0.001**				**<0.001**			
95% CI	**[58.97, 70.46]**				**[5.36, 9.06]**				**[0.28, 0.97]**			
Cohen's d	**1.10**				**0.38**				**0.17**			

Table 10. Summary of main findings of the augmentation effect of collaborating with ChatGPT for Study 4, Task 1 performance. *DV* = dependent variable, SD = standard deviation, CI = confidence interval, Cohen's d is the standardized effect size. Comparisons were conducted between participants who worked solo in Task 1 (i.e., the Solo-Solo and Solo-Collab conditions) and participants who collaborated with AI in Task 1 (i.e., the Collab-Solo and Collab-Collab conditions). T-tests were used to evaluate mean differences for each dependent variable. Significant results are bolded for emphasis.

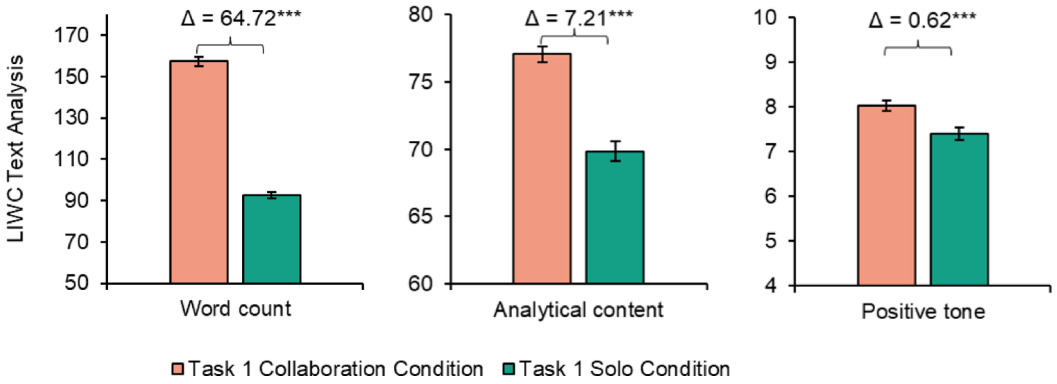


Fig. 12. Performance augmentation of ChatGPT of Study 4, Task 1. We analyzed Task 1 performance using LIWC across three dimensions: Word count, analytical content, and positive tone. Results indicated that collaboration with ChatGPT (i.e., the Collab-Solo and Collab-Collab conditions, $N = 801$) enhanced Task 1 performance significantly, with participants who used ChatGPT producing outputs that were longer, more analytical, and more positive in tone compared to those who worked solo in Task 2 (i.e., the Solo-Solo and Solo-Collab conditions, $N = 823$). $^{*}p < .05$, $^{***}p < .001$. Error bars indicate ± 1 SEM.

collaboration with GenAI does not significantly improve the subsequent solo task performance. (see Table 11; Fig. 13 for the results summary).

Deprivation effect of human-GenAI collaboration on psychological experiences (RQ2-4)
Finally, we examined changes in psychological experience measures from Task 1 to Task 2 across four different conditions. Participants reported their sense of control, intrinsic motivation, and feelings of boredom after completing Task 1 and Task 2, respectively. We conducted mixed-design repeated measures ANOVAs to analyze the interaction effects between task (Task 1 vs. Task 2, a within-subjects factor) and condition (described in detail below, a between-subjects factor). Additionally, paired-samples *t*-tests were conducted to assess within-individual changes in psychological experiences from Task 1 to Task 2 within each condition. The key findings from both the ANOVAs and *t*-tests are summarized in Table 12 and illustrated in Fig. 14.

Change of sense of control with task transition Results of paired-samples *t*-tests indicated that participants reported different levels of sense of control from Task 1 to Task 2. For the Solo-Solo condition, participants reported a significant decrease in their sense of control, from Task 1 ($M = 5.99$, $SD = 1.00$) to Task 2 ($M = 5.84$, $SD = 1.05$), $t(427) = 3.69$, $p < .001$, 95% $CI = [0.07, 0.24]$, $d = 0.15$. For the Collab-Solo condition, participants reported a significant increase in their sense of control from Task 1 ($M = 5.70$, $SD = 1.15$) to Task 2 ($M = 6.12$, $SD = 1.00$), $t(344) = -6.22$, $p < .001$, 95% $CI = [-0.55, -0.29]$, $d = -0.39$. For the Solo-Collab condition, participants experienced a significant decrease in the sense of control from Task 1 ($M = 6.01$, $SD = 0.95$) to Task 2 ($M = 5.00$, $SD = 1.41$), $t(394) = 13.23$, $p < .001$, 95% $CI = [0.86, 1.17]$, $d = 0.84$. For the Collab-Collab condition, participants experienced a similar level of sense of control from Task 1 ($M = 5.74$, $SD = 1.11$) to Task 2 ($M = 5.69$, $SD = 1.12$), $t(455) = 0.181$, $p = .181$, 95% $CI = [-0.03, 0.13]$, $d = 0.04$.

We further conducted ANOVAs to compare changes in participants' sense of control from Task 1 to Task 2 across conditions. A significant interaction effect between task and condition was observed when comparing the Solo-Solo and Collab-Solo conditions, $F(1, 771) = 56.36$, $p < .001$, partial $\eta^2 = 0.068$, $\omega^2 = 0.067$. This medium effect size indicates a meaningful boost in perceived control during the transition to solo work, consistent with our previous studies. Specifically, participants who initially collaborated with GenAI experienced a significant increase in their sense of control when transitioning to solo work ($\Delta = 0.42$, $p < .001$), whereas those who worked solo throughout reported a gradual decline ($\Delta = -0.15$, $p < .001$). This pattern suggests that while collaboration with GenAI may temporarily reduce perceived control, individuals regain it when working independently, whereas sustained solo work may lead to a subtle erosion of control over time.

When comparing the Solo-Collab and Solo-Solo conditions, another significant interaction effect was observed, $F(1, 738) = 101.02$, $p < .001$, partial $\eta^2 = 0.110$, $\omega^2 = 0.108$. Participants who initially worked solo and later collaborated with GenAI reported a sharp decline in their sense of control ($\Delta = -1.01$, $p < .001$), whereas those who worked solo throughout experienced a more modest decrease ($\Delta = -0.15$, $p < .001$). The substantial drop in the Solo-Collab condition suggests that transitioning from solo work to GenAI collaboration significantly undermines perceived sense of control, more so than sustained solo work alone. This finding highlights the psychological cost of shifting from independent work to AI-assisted collaboration, reinforcing the notion that the presence of GenAI may erode an individual's autonomy over their work.

Next, we compared the Collab-Solo condition with the other two conditions. When comparing the Collab-Solo and Solo-Collab conditions, a significant interaction effect was found, $F(1, 738) = 191.91$, $p < .001$, partial $\eta^2 = 0.206$, $\omega^2 = 0.205$. Participants who initially collaborated with GenAI and subsequently worked solo reported an increase in their sense of control. In contrast, those who started solo and then collaborated with GenAI exhibited a sharp decline in their sense of control ($\Delta = -1.01$, $p < .001$). This substantial drop underscores the

DV	Word Count				Analytical Content				Positive Tone			
Condition	Solo-Solo	Collab-Solo	Solo-Collab	Collab-Collab	Solo-Solo	Collab-Solo	Solo-Collab	Collab-Collab	Solo-Solo	Collab-Solo	Solo-Collab	Collab-Collab
Mean	88.47	81.77	162.49	179.97	72.21	74.32	77.67	80.34	7.53	7.10	7.82	7.43
SD	50.02	42.55	68.19	62.65	21.68	21.27	15.58	12.66	4.18	3.73	3.32	2.90
t (df)	**-1.98 (771)**				1.35 (771)				-1.53 (762.94)			
p	**0.048**				0.177				0.128			
95% CI	[**-13.35**, **-0.05**]				[-0.95, 5.16]				[-1.00, 0.12]			
Cohen's d	**0.14**				0.10				0.11			

Table 11. Summary of main findings of Study 4, Task 2 performance across conditions. *DV* = dependent variable, SD = standard deviation, CI = confidence interval, Cohen's d is the standardized effect size. Comparisons were conducted between Collab-Solo (transitioning from collaboration to solo work) and Solo-Solo (working solo in both tasks) conditions. T-tests were used to evaluate mean differences for each dependent variable. Significant results are bolded for emphasis.

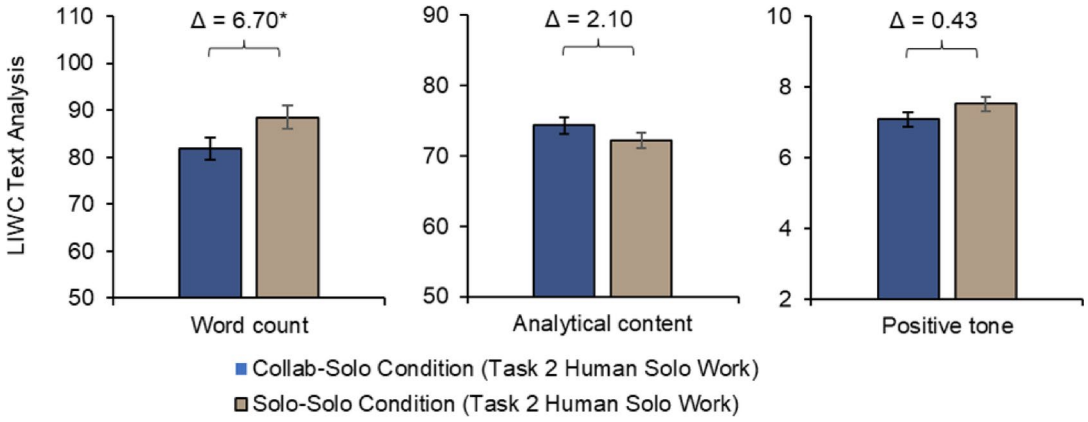


Fig. 13. Study 4, Task 3 performance across Collab-Solo and Solo-Solo conditions. Task 2 performance was evaluated based on text length, analytical content, and positive tone of participants' outputs, analyzed using the LIWC software. Results showed no significant differences between the two conditions in analytical content or positive tone, but a slight difference in text length, with the Solo-Solo condition producing longer texts. These findings suggest that prior collaboration with ChatGPT does not significantly enhance performance in subsequent solo tasks. $N = 773$. $^{*}p < .05$. Error bars indicate ± 1 SEM.

potential psychological cost of transitioning from independent work to collaboration with GenAI, suggesting that working with GenAI can undermine human workers' sense of control, regardless of task sequence.

When comparing the Collab-Solo and Collab-Collab conditions, another significant interaction effect was observed, $F(1, 799) = 40.02$, $p < .001$, partial $\eta^2 = 0.048$, $\omega^2 = 0.046$. Participants who initially collaborated with GenAI and subsequently worked solo reported a significant increase in their sense of control, whereas those who collaborated with GenAI consistently across both tasks reported a relatively stable sense of control, with only a slight decrease (Δ = -0.05, $p = .181$). These findings suggest that while prolonged collaboration with GenAI maintains a consistent level of perceived control, it does not enhance it. Moreover, the observed increase in the Collab-Solo condition indicates that individuals regain a sense of control once they transition to independent work. This pattern further supports our argument that working with GenAI can initially diminish human workers' sense of control, but this effect can be mitigated when they return to solo work.

Change of intrinsic motivation with task transition Results from paired-samples *t*-tests indicated that participants consistently reported a significant decrease in intrinsic motivation from Task 1 to Task 2 across all four conditions. Specifically, for the Solo-Solo condition, intrinsic motivation declined from Task 1 ($M = 4.99$, $SD = 1.34$) to Task 2 ($M = 4.79$, $SD = 1.46$), $t(427) = 3.39$, $p = .001$, 95% CI = [0.08, 0.31], $d = 0.14$. For the Collab-Solo condition, participants reported a steeper decline in intrinsic motivation from Task 1 ($M = 5.06$, $SD = 1.34$) to Task 2 ($M = 4.65$, $SD = 1.52$), $t(344) = 5.47$, $p < .001$, 95% CI = [0.27, 0.56], $d = 0.29$. For the Solo-Collab condition, a similar decline was observed, with intrinsic motivation decreasing from Task 1 ($M = 5.09$, $SD = 1.32$) to Task 2 ($M = 4.49$, $SD = 1.52$), $t(394) = 7.72$, $p < .001$, 95% CI = [0.44, 0.74], $d = 0.42$. For the Collab-Collab condition, intrinsic motivation also decreased from Task 1 ($M = 5.16$, $SD = 1.31$) to Task 2 ($M = 4.88$, $SD = 1.44$), $t(455) = 6.00$, $p < .001$, 95% CI = [0.19, 0.37], $d = 0.20$.

We further conducted ANOVAs to compare changes in intrinsic motivation from Task 1 to Task 2 across conditions. A significant interaction effect between task and condition was observed when comparing the Solo-Solo and Collab-Solo conditions, $F(1, 771) = 5.45$, $p = .020$, partial $\eta^2 = 0.007$, $\omega^2 = 0.006$. Participants in the Collab-Solo condition experienced a greater decline (Δ = -0.41, $p < .001$) in intrinsic motivation compared to those in the Solo-Solo condition (Δ = -0.20, $p = .001$). These results align with prior findings, highlighting the motivational challenges workers face when transitioning from AI-assisted collaboration to solo tasks.

When comparing the Solo-Collab and Solo-Solo conditions, a significant interaction effect was observed, $F(1, 821) = 17.36$, $p < .001$, partial $\eta^2 = 0.021$, $\omega^2 = 0.019$. Participants who initially worked solo and then transitioned to GenAI collaboration experienced a significantly larger decline in intrinsic motivation (Δ = -0.60, $p < .001$) compared to those who worked solo throughout (Δ = -0.20, $p = .001$). This notable difference suggests that the shift from independent work to AI-assisted collaboration exacerbates motivational decline more than sustained solo work. These findings indicate that despite the potential efficiency benefits of collaborating with GenAI, it does not necessarily enhance or sustain intrinsic motivation and may even accelerate its erosion when introduced after an initial period of independent work.

Next, we compared the Collab-Solo condition with the Solo-Collab condition. While intrinsic motivation declined in both conditions, the decrease was larger in the Solo-Collab condition (Δ = -0.60, $p < .001$) compared to the Collab-Solo condition (Δ = -0.41, $p < .001$). However, this difference did not reach statistical significance, $F(1, 738) = 2.70$, $p = .101$, partial $\eta^2 = 0.004$, $\omega^2 = 0.002$. This suggests that although the transition sequences differ (solo to collaboration vs. collaboration to solo), there is no strong statistical evidence indicating that one transition is more detrimental to intrinsic motivation than the other. While Solo-Collab shows a larger absolute decline, the variability in participants' responses means the difference is not statistically robust. Thus, both types

DV	Sense of Control							
Condition	Solo-Solo (1)		Collab-Solo (2)		Solo-Collab (3)		Collab-Collab (4)	
Task	Task 1	Task 2	Task 1	Task 2	Task 1	Task 2	Task 1	Task 2
Mean	5.99	5.84	5.70	6.12	6.01	5.00	5.74	5.69
SD	1.00	1.05	1.15	1.00	0.95	1.41	1.11	1.12
t (*df*)	**3.69 (427)**		**-6.22 (344)**		**13.23 (394)**		1.34 (455)	
p	**< 0.001**		**< 0.001**		**< 0.001**		0.181	
95% *CI*	**[0.07, 0.24]**		**[-0.55, -0.28]**		**[0.86, 1.17]**		[-0.03, 0.13]	
Cohen's *d*	**0.15**		**-0.39**		**0.84**		0.04	
	F test (*Cond 2 vs. 1*)		F test (*Cond 3 vs. 2*)		F test (*Cond 3 vs. 1*)		F test (*Cond 4 vs. 2*)	
F (df1, df2)	**56.36 (1, 771)**		**191.91 (1, 738)**		**101.02 (1, 738)**		**40.02 (1, 799)**	
p	**< 0.001**		**< 0.001**		**< 0.001**		**< 0.001**	
partial η^2	**0.068**		**0.206**		**0.110**		**0.048**	
ω^2	**0.067**		**0.205**		**0.108**		**0.046**	
DV	**Intrinsic Motivation**							
Condition	**Solo-Solo (1)**		**Collab-Solo (2)**		**Solo-Collab (3)**		**Collab-Collab (4)**	
Task	Task 1	Task 2	Task 1	Task 2	Task 1	Task 2	Task 1	Task 2
Mean	4.99	4.79	5.06	4.65	5.09	4.49	5.16	4.88
SD	1.34	1.46	1.34	1.52	1.32	1.52	1.31	1.44
t (*df*)	**3.39 (427)**		**5.47 (344)**		**7.72 (394)**		**6.00 (455)**	
p	**0.001**		**< 0.001**		**< 0.001**		**< 0.001**	
95% *CI*	**[0.08, 0.31]**		**[0.27, 0.56]**		**[0.44, 0.74]**		**[0.19, 0.37]**	
Cohen's *d*	**0.14**		**0.29**		**0.42**		**0.20**	
	F test (*Cond 2 vs. 1*)		F test (*Cond 3 vs. 2*)		F test (*Cond 3 vs. 1*)		F test (*Cond 4 vs. 2*)	
F (df1, df2)	**5.45 (1, 771)**		2.70 (1, 738)		**17.36 (1, 821)**		2.53 (1, 799)	
p	**0.020**		0.101		**< 0.001**		0.112	
partial η^2	**0.007**		0.004		**0.021**		0.003	
ω^2	**0.006**		0.002		**0.019**		0.002	
DV	**Feelings of Boredom**							
Condition	**Solo-Solo (1)**		**Collab-Solo (2)**		**Solo-Collab (3)**		**Collab-Collab (4)**	
Task	Task 1	Task 2	Task 1	Task 2	Task 1	Task 2	Task 1	Task 2
Mean	2.68	3.00	2.61	3.05	2.57	3.07	2.50	2.75
SD	1.25	1.45	1.28	1.47	1.19	1.38	1.23	1.44
t (*df*)	**-5.44 (427)**		**-5.85 (344)**		**-6.71 (394)**		**-4.88 (455)**	
p	**< 0.001**		**< 0.001**		**< 0.001**		**< 0.001**	
95% *CI*	**[-0.43, -0.20]**		**[-0.59, -0.29]**		**[-0.65, -0.35]**		**[-0.35, -0.15]**	
Cohen's *d*	**-0.24**		**-0.32**		**-0.39**		**-0.19**	
	F test (*Cond 2 vs. 1*)		F test (*Cond 3 vs. 2*)		F test (*Cond 3 vs. 1*)		F test (*Cond 4 vs. 2*)	
F (df1, df2)	1.72 (1, 771)		0.36 (1, 738)		**3.92 (1, 821)**		**4.85 (1, 799)**	
p	0.190		0.551		**0.048**		**0.028**	
partial η^2	0.002		0.000		**0.005**		**0.006**	
ω^2	0.001		-0.001		**0.004**		**0.005**	

Table 12. Summary of main findings of psychological experiences of Study 4. *DV* = dependent variable, *SD* = standard deviation, *CI* = confidence interval, Cohen's *d* is the standardized effect size, partial η^2 and ω^2 indicate effect sizes in repeated measures ANOVAs. Paired-samples *t*-tests were conducted to assess within-individual cross-task differences, while repeated measures ANOVAs evaluated interaction effects. All significant results are bolded for emphasis.

of transitions—starting with AI collaboration and then working solo, or vice versa—can negatively impact intrinsic motivation to a similar extent.

Similarly, for the comparison between the Collab-Solo and Collab-Collab conditions, no significant interaction effect was observed, $F(1, 799) = 2.53$, $p = .112$, partial $\eta^2 = 0.003$, $\omega^2 = 0.002$. This suggests that continuing to collaborate with GenAI does not provide a distinct advantage over transitioning from GenAI-

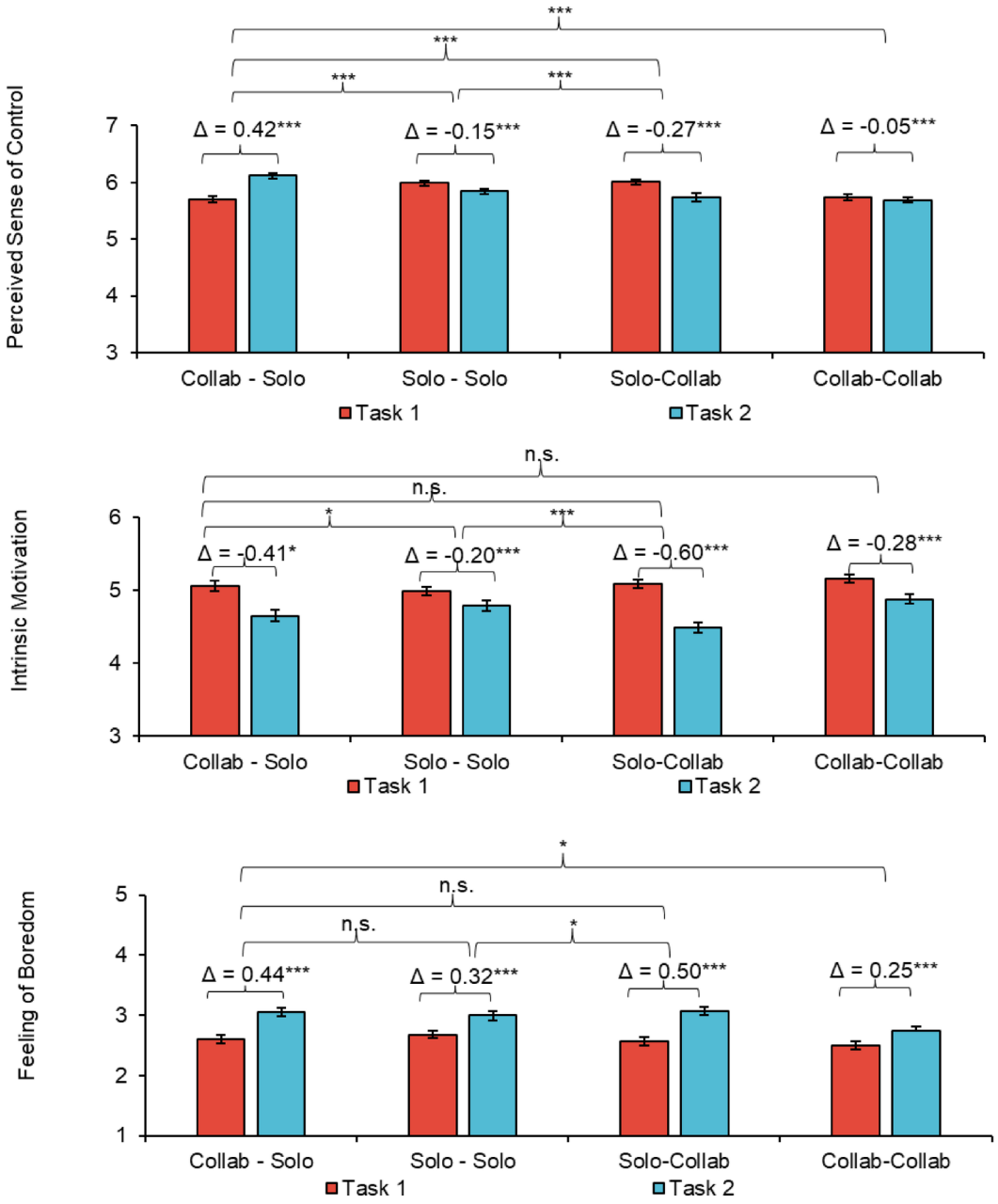


Fig. 14. Study 4, Changes of psychological experiences from Task 1 to Task 2 across four conditions. Participants rated their sense of control, intrinsic motivation, and feelings of boredom on 7-point scales after completing each task. Those who transitioned from collaboration with ChatGPT to solo work ($N = 345$) reported an increased sense of control, decreased intrinsic motivation, and increased feelings of boredom. Participants who worked solo in both tasks ($N = 428$) experienced a decrease in sense of control and intrinsic motivation, and an increase in feelings of boredom, similar to those in the Solo-Collab condition ($N = 395$), who transitioned from solo work to collaboration with ChatGPT, and those in the Collab-Collab condition ($N = 456$), who collaborated with ChatGPT in both tasks. Comparisons between the Collab-Solo condition and the other three conditions are illustrated in the figure. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$. Error bars indicate ± 1 SEM.

assisted collaboration to solo work in maintaining intrinsic motivation. In other words, remaining in a GenAI-supported collaborative environment does not necessarily buffer against the decline in intrinsic motivation seen across tasks.

Change of feelings of boredom with task transition Results from paired-samples t-tests indicated that participants consistently reported a significant increase in feelings of boredom from Task 1 to Task 2 across all four conditions. Specifically, for the Solo-Solo condition, feelings of boredom significantly increased from Task 1 ($M = 2.68$, $SD = 1.25$) to Task 2 ($M = 3.00$, $SD = 1.45$), $t(427) = -5.44$, $p < .001$, 95% $CI = [-0.43, -0.20]$, $d = -0.24$.

For the Collab-Solo condition, participants reported a significant increase in feelings of boredom from Task 1 ($M = 2.61$, $SD = 1.28$) to Task 2 ($M = 3.05$, $SD = 1.47$), $t(344) = -5.85$, $p < .001$, 95% $CI = [-0.59, -0.29]$, $d = -0.32$. For the Solo-Collab condition, a significant increase in boredom was also observed, from Task 1 ($M = 2.57$, $SD = 1.19$) to Task 2 ($M = 3.07$, $SD = 1.38$), $t(394) = -6.71$, $p < .001$, 95% $CI = [-0.65, -0.35]$, $d = -0.39$. For the Collab-Collab condition, participants also reported an increase in feelings of boredom from Task 1 ($M = 2.50$, $SD = 1.23$) to Task 2 ($M = 2.75$, $SD = 1.44$), $t(455) = -4.88$, $p < .001$, 95% $CI = [-0.35, -0.15]$, $d = -0.19$.

We further conducted ANOVAs to compare changes in participants' feelings of boredom from Task 1 to Task 2 across conditions. When comparing the Solo-Solo and Collab-Solo conditions, the analysis did not show a significant interaction effect between task and condition, $F(1, 771) = 1.72$, $p = .190$, partial $\eta^2 = 0.002$, $\omega^2 = 0.001$. This suggests that transitioning from GenAI-assisted collaboration to solo work did not lead to a significantly different change in boredom compared to working solo across both tasks ($\Delta = 0.44$, $p < .001$ vs. $\Delta = 0.32$, $p < .001$, respectively). In both conditions, participants experienced a similar increase in boredom.

When comparing the Solo-Collab and Solo-Solo conditions, a significant interaction effect was observed, $F(1, 821) = 3.92$, $p = .048$, partial $\eta^2 = 0.005$, $\omega^2 = 0.004$. Participants who initially worked solo and later collaborated with GenAI experienced a slightly greater increase in boredom ($\Delta = 0.50$, $p < .001$) compared to those who worked solo throughout ($\Delta = 0.32$, $p < .001$). This difference, though statistically significant, suggests that the transition from solo to GenAI-assisted collaboration may introduce additional cognitive or psychological strain, contributing to a steeper rise in boredom than sustained solo work. However, given the small effect size, this impact is modest. These findings indicate that while boredom increases across all conditions, shifting from independent work to AI collaboration does not necessarily alleviate it and may, in some cases, intensify the experience.

Next, when comparing the Collab-Solo condition with the Solo-Collab condition, no significant interaction effect was observed, $F(1, 738) = 0.36$, $p = .551$, partial $\eta^2 = 0.000$, $\omega^2 = -0.001$). This indicates that the sequence of solo and collaboration work (whether starting solo and transitioning to GenAI-assisted collaboration or vice versa) did not have a meaningful impact on changes in boredom. Regardless of the order of collaboration, participants reported a comparable increase in boredom.

Finally, the comparison between the Collab-Solo condition and the Collab-Collab condition revealed a significant interaction effect, $F(1, 799) = 4.85$, $p = .028$, partial $\eta^2 = 0.006$, $\omega^2 = 0.005$. Participants in the Collab-Solo condition reported a greater increase in boredom compared to those who worked with GenAI throughout (Collab-Collab). This suggests that while collaboration with GenAI does not eliminate boredom entirely, prolonged collaboration may help mitigate its increase over time. In contrast, transitioning from collaboration to solo work may exacerbate feelings of boredom, potentially due to the absence of external assistance and engagement previously provided by GenAI.

Navigating the human-AI work paradigm: insights and implications

There has been considerable discussion recently about the skills and abilities individuals should develop to avoid being replaced by AI and to maintain their unique contributions[42]. However, another important question is whether this increasingly common "dance with AI" work paradigm can sustain our intrinsic motivation in continuous work endeavors. Our work investigated the potential deprivation effects of AI collaboration on human motivation, contributing to the discourse on the overlooked negative aspects of human-AI interactions. From the perspective of well-being and sustained engagement in work, intrinsic motivation, a sense of control, and the avoidance of boredom are essential psychological experiences that not only enhance productivity but also contribute to long-term job satisfaction and cannot be fully compensated by extrinsic rewards such as monetary incentives.

Across four studies involving various cognitively demanding tasks (i.e., text generation and creative brainstorming), our results consistently demonstrate a double-edged sword effect of human-GenAI collaboration. While GenAI enhances the quality of human-generated outputs, it fails to sustain subsequent task performance and undermines key psychological experiences, including sense of control, intrinsic motivation, and feeling of boredom. While GenAI contributes substantially to the current task outcomes, individuals may perceive a reduction in personal agency, thereby undermining their sense of control. Furthermore, the transition from engaging, AI-assisted tasks to less stimulating, human-solo tasks may exacerbate feelings of boredom, as tasks lacking novelty and challenges are known to erode sustained motivation[34].

In particular, Study 4 provides a deeper understanding of the psychological deprivation effect in human-GenAI collaboration by introducing two additional experimental conditions—Solo-Collab and Collab-Collab—alongside the previously examined Solo-Solo and Collab-Solo conditions. First, the findings highlight the important role of task sequence. Transitioning from collaboration to solo work partially restores the sense of control, whereas transitioning from solo work to collaboration diminishes it. This suggests that the order in which workers engage in GenAI-assisted collaboration and solo tasks shapes their psychological experience. Second, prolonged collaboration with GenAI has a dual effect. While it helps maintain a relatively stable sense of control, it does not enhance intrinsic motivation or alleviate the increase in boredom. This reinforces the psychological costs of sustained human-GenAI collaboration. Lastly, task transitions—whether from collaboration to solo work or vice versa—negatively impact intrinsic motivation and sense of control. This further supports the existence of a psychological deprivation effect in human-GenAI collaboration, emphasizing the challenges workers face when alternating between these modes of work.

Contrary to our expectations, four studies provided little evidence of performance augmentation spillover effects from GenAI. This lack of spillover may stem from negative psychological experiences that overshadowed GenAI's potential benefits, limiting its ability to enhance subsequent task performance. Notably, participants transitioning from GenAI collaboration to solo work consistently reported psychological deprivation effects, including reduced intrinsic motivation and increased boredom, despite a heightened sense of control. These

Research Focus	Performance Augmentation	Spillover of Performance Augmentation	Psychological Deprivation
Analysis	Compare Task 1 performance between the Collab-Solo and Solo-Solo conditions	Compare Task 2 performance between the Collab-Solo and Solo-Solo conditions	Compare changes of psychological experiences from Task 1 to Task 2 between the Collab-Solo and Solo-Solo conditions
Study 1 Findings	Collaboration with GenAI enhanced the quality of promotional posts compared to work by human participants alone.	GenAI had a performance spillover effect on the idea creativity, but not on the idea quantity in the subsequent alternative uses test task.	Transitioning from working with GenAI to solo work increased sense of control, decreased intrinsic motivation, and increased boredom.
Study 2 Findings	Performance review reports co-created with GenAI were longer, more analytical and more prosocial compared to reports created by human participants alone.	GenAI had NO performance spillover effects on the quantity, novelty, or usefulness of ideas in the subsequent product improvement task.	Transitioning from working with GenAI to solo work increased sense of control, decreased intrinsic motivation (n.s.), and increased boredom (marginally significant).
Study 3 Findings	Emails co-created with GenAI were longer, more affiliative and socially expressive compared to those drafted by human participants alone.	GenAI had NO performance spillover effects in the quantity, novelty, or usefulness of ideas in the subsequent product promotion task.	Transitioning from working with GenAI to solo work increased sense of control, decreased intrinsic motivation, and increased boredom.
Study 4 Findings	Collaboration with GenAI led to better performance in the text-generation task (longer, more analytical, and more positive in tone).	GenAI had NO performance spillover effects on the text length, analytical content, or positive tone in the subsequent solo task.	Transitioning from working with GenAI to solo work increased sense of control, decreased intrinsic motivation, and increased boredom (n.s.).

Table 13. Summary of main findings across studies.

effects were absent or less pronounced in participants who completed both tasks independently. According to self-determination theory, such declines in intrinsic motivation can undermine sustained engagement and performance. Study 4 further eliminated the possibility that the absence of spillover effects was due to differences in task types or display sequence. This was achieved by using two similar text-generation tasks and counterbalancing their display order. These findings suggest that the disruption caused by transitioning from GenAI collaboration to solo work likely hinders the potential for performance spillover.

Our findings extend the growing body of literature on human-GenAI interaction by highlighting its broader psychological implications. While prior research has emphasized the productivity benefits of GenAI[3,4], our study shifts focus to the longer-term sustainability of work engagement in AI-augmented environments. This approach offers a unique perspective on the psychological trade-offs in AI partnerships, which resonates with broader debates on the societal impact of AI. Second, our research emphasizes the sustainable development of human workers by examining their psychological experiences in tasks following AI collaboration. This shift in focus highlights not just the immediate benefits of AI in enhancing performance but also the ongoing and longer-term impacts on worker well-being. Our research marks an initial exploration into how task transitions involving GenAI and human solo work influence psychological experiences. A summary of findings across the four studies is presented in Table 13.

We acknowledge that our study has certain limitations. First, our research design is limited to only two consecutive tasks. We were not able to investigate whether the augmentation and deprivation effect persists in scenarios involving multiple tasks. Moreover, our participants were sourced from an online platform, not reflecting real-world workplace dynamics of human-AI collaboration. The AI collaboration in real work settings might diverge from our experimental simulations. Additionally, the incentives offered to participants could introduce extrinsic motivators, potentially skewing their psychological experiences and task outcomes. Finally, we did not examine the potential curvilinear effects of intrinsic motivation in this paper. Previous research suggests that intrinsic motivation may exert a curvilinear effect—moderate levels of intrinsic motivation can carry over to other tasks, but extremely high levels may create contrasting effects and reduce enjoyment of subsequent task[34,43]. However, across four studies our participants' intrinsic motivation levels in Task 1 were relatively high but not extreme (around 5 on a 7-point scale). Thus, we did not anticipate or observe nonlinear effects in our study. Future research could investigate the conditions under which curvilinear effects of intrinsic motivation emerge, particularly in tasks with varying levels of complexity, interest, or cognitive demand.

Nevertheless, our findings bear insightful implications for professionals and policymakers. For example, AI system designers should emphasize human agency in collaborative platforms, achieved by integrating user feedback, input, and customization, ensuring users retain a sense of control during collaborations with AI. In addition, our paper also has implications for job designers, emphasizing the significance of maintaining individual psychological well-being while at the same time reaping the benefits of AI technologies. This understanding can help create fulfilling and engaging tasks, promoting sustained employee motivation and meaningful collaboration by aligning tasks with workers' preferences and skills. Finally, human workers can actively craft their jobs between collaboration with AI and independent work, promoting a sense of fulfillment and motivation in their professional endeavors.

Methods

We conducted four pre-registered online experiments, all designed and administrated on the Qualtrics survey platform. Participants provided informed consent via the Qualtrics survey platform panel. Participants were sourced from Prolific, a crowdsourcing survey platform frequently employed by scholars in management and social psychology. The experiments comply with all relevant ethical regulations and were approved for human subject participation by the Institutional Review Board by the corresponding author's affiliated institution (Approval no. 20230323). We have transparently reported all conditions, measures, and exclusions.

Analytical strategies

To investigate the augmentation effect of human-GenAI collaboration, we employed independent sample *t*-tests on the performance of the first task across both groups. Similarly, for testing the spillover effect of AI augmentation, we applied independent sample *t*-tests to examine the differences between participants' performance in the second task. To examine changes in individuals' psychological experiences across the two tasks, we employed three sperate 2×2 mixed-design repeated measures ANOVAs and paired-samples *t*-tests to investigate the effects of human-GenAI collaboration on participants' sense of control, intrinsic motivation, and feelings of boredom across two distinct tasks, respectively.

Study 1

We based our sample size on a priori power analyses conducted with G*Power[44,45]. The analysis (with $1 - \beta = 0.80$, $\alpha = 0.05$) suggested a sample of 352 participants to detect a small to medium effect of $d = 0.30$.

Participants and procedures

Leveraging the prescreen tool integrated into Prolific, we specifically targeted participants who were native English speakers residing in the UK and had an approval rating exceeding 95%. We collected responses from 352 participants, with none being excluded from the study. Of participants, 50.0% identified as women. Approximately 45.5% of respondents held an associate degree or higher. The participants' average age was 41.02 years ($SD = 14.27$), and 73.6% were employed. Upon successful survey completion, participants were compensated with 2.25 British pounds.

Task 1: crafting a promotional Facebook post

In Task 1, participants undertook a promotional post-writing assignment. They were asked to imagine themselves as marketing specialists at a startup company which is about to launch an eco-friendly reusable water bottle named "HydraGreen" in the coming month. Their objective was to craft a promotional Facebook post for this product, spotlighting its standout features: a built-in filter, a convenient carry handle with a carabiner clip, a contemporary design available in diverse colors and sizes, and robust construction. The post should attract potential customers and increase their purchase intentions.

Participants were randomly assigned to one of two conditions: human-ChatGPT collaboration ($N = 176$) or human working-alone ($N = 176$). Those in the human working-alone condition completed the writing task independently. Participants in the human-ChatGPT collaboration condition were instructed to craft the post with the assistance of ChatGPT version 3.5. This integration was achieved using Qualtrics' Web Service function, allowing for seamless user experience within the survey interface. A detailed description of the integration process can be found in the supplementary materials. This integration approach was inspired by the methods outlined previously[46]. Prior to working on the post, participants were briefed on ChatGPT's functionality and practiced by texting prompts. This allowed them to familiarize themselves with ChatGPT and ensured they recognized they were interacting with genuine GenAI. After this hands-on session, the writing task was reintroduced, accompanied by a prompt input space. Participants then received a response generated by ChatGPT and were asked to complete the post based on its response. Following task completion, both groups answered a series of survey questions designed to measure their psychological experiences during the task and assess the effectiveness of our manipulation.

Task 2: the alternative uses test

Upon completing Task 1, participants from both conditions were presented with the Alternative Uses Test (AUT)[40]. In this task, they were instructed to brainstorm as many innovative uses for a soda can as possible and to type their ideas in a text box in three minutes. After this, participants answered another series of questions designed to measure their psychological experiences during the AUT and collect their demographic information.

Post-task questionnaires

We measured participants' perceived sense of control, intrinsic motivation, and feelings of boredom after both Task 1 and Task 2, respectively. Beyond these measures, demographic information such as age, gender, education, and employment status were collected at the end of the questionnaire.

Manipulation check

Participants completed a manipulation check after submitting their writing posts in Task 1. One question asked them to indicate how they completed Task 1, choosing between "I created the Facebook post purely on my own" and "I collaborated with an artificial intelligence (AI) tool in creating the Facebook post". A chi-square test of independence revealed a significant relationship between condition and reported collaboration method (χ^2 (1) = 218.43, $p < .001$). 81.3% of participants in the human-Al collaboration condition reported collaborating with AI compared to only 18.8% in the human-working-alone condition. Only 3.4% of participants in the human-working-alone condition reported collaborating with AI compared to 96.6% in the human-working-alone condition. This suggests the manipulation of collaboration method was successful.

Sense of control

Adapting from previous work[47], we measured participants' sense of control after each task on a seven-point scale with three items (from 1 = "strongly disagree" to 7 = "strongly agree"). An example item is "I felt in control of the task". The reliability of the measure suggested good consistency, with Cronbach's α values of 0.88 in Task 1 and 0.89 in Task 2.

Intrinsic motivation
We used a four-item measure from previous work[34] to assess the extent to which participants found the task interesting, enjoyable, fun, and engaging (from 1 = "strongly disagree to" 7 = "strongly agree"). The measure demonstrated excellent consistency, with Cronbach's α values of 0.96 in Task 1 and 0.95 in Task 2.

Boredom
Following previous work[34], we assessed perceived task boredom using four items[39]. Participants were asked to indicate how they felt while performing the task (from 1 = "strongly disagree" to 7 = "strongly agree"), with statements such as "I thought that the task served no important purpose". The Cronbach's α is 0.41 for Task 1 and 0.51 for Task 2, indicating low internal consistency. We acknowledge the limitations of the low reliability of this boredom scale, and results involving this scale should be interpreted cautiously.

Assessment of Task 1 performance
Performance for Task 1 was assessed by five independent raters who evaluated the creativity of participants' inputs following a well-established approach[22]. To ensure a standardized evaluation, comprehensive task-specific instructions were developed to establish clear evaluation criteria. These manuals incorporated information from preceding tasks, detailed explanations of the evaluation dimensions, and representative examples for rating. Five undergraduate students proficient in English reading were recruited for this task, each meeting the criterion of a TOEFL reading subscore of 25 or above, or an IELTS reading subscore of 8 or above. Comprehensive briefings on the respective instruction manuals were given to ensure clarity and understanding.

The Facebook posts were assessed based on two primary dimensions: "engaging" and "informative" on a five-point scale (from 1 = "poor" to 5 = "excellent"). The "engaging" dimension refers to the extent to which the post captivates readers, elicits positive reactions, and fosters purchasing intentions. This could be achieved through techniques such as weaving in real-life scenarios or adopting a spirited tone. The "informative" dimension evaluated the efficacy of the post in conveying pertinent product details, which included emphasizing key features and providing a thorough overview of the product's functionalities. We computed intraclass correlation coefficients (ICCs) using a two-way random effects model and a consistency definition[48,49]. The ICC was 0.87 for ratings of "engaging" and 0.86 for ratings of "informative", indicating good interrater reliability.

Assessment of Task 2 performance
For Task 2, the alternative uses provided by participants were evaluated by five independent raters following the Open Creativity Scoring method as detailed on the website: https://openscoring.du.edu/. The alternative uses were assessed based on two dimensions: idea quantity (i.e., the number of ideas) and idea quality (i.e., the extent to which ideas reflect creative thinking).

For idea quantity, one research assistant who was blind to the study counted the number of ideas that each participant generated, which was then cross verified by a second research assistant. Idea quality refers to the degree to which the proposed use was innovative, original, imaginative, and clever. A truly creative alternative use would extend beyond the traditional applications of a soda can. Following the same procedures from Task 1, we recruited another five undergraduate students who were proficient in English and provided them with detailed guidelines to ensure assessment accuracy and uniformity. Those raters assessed each participant's set of ideas (i.e., the overall creativity of all ideas) on a five-point scale (from 1 = "poor" to 5 = "excellent"). The intraclass correlation coefficient (ICC) using a two-way random effects model and a consistency definition was 0.67, suggesting moderately good interrater reliability[48,49].

Study 2

The objective of Study 2 is to undertake a high-powered replication of Study 1's findings, employing varied task scenarios to bolster the findings' robustness and generalizability. In this study, we introduced two distinct task scenarios: "Drafting a Performance Review Report" for Task 1 and "Brainstorming for Product Improvement" for Task 2. Following the guidance of previous work[50,51], we recalibrated our power analysis to "$d = 0.20$" as the smallest effect size of interests (SESOI) for our study. This adjustment ensures our study is sufficiently powered to identify subtle yet theoretically important effects, aligning our methodology with the latest standards in research design for human-AI interaction studies. Therefore, we conducted a priori power analysis using G*Power, factoring in this effect size. To detect this effect with an α error probability of 0.05 and a power (1-β error probability) of 0.80, the estimated sample size required for our replication study is 788.

Participants and procedures
We targeted participants who were native English speakers residing in the UK and had an approval rating exceeding 95%. Meanwhile, to capture a representative sample of working adults within a broad age range, ensuring relevance to the professional interactions between human workers and AI, we prescreened those being employed either part-time or full-time within the age range of 18–60. We collected responses from 793 participants, with none being excluded from the study. Of participants, 49.9% identified as women ($N = 396$). Approximately 74.15% of respondents held an associate degree or higher ($N = 588$). The participants' average age was 36.18 years ($SD = 10.22$). Upon successful survey completion, participants were compensated with 2.25 British pounds.

Task 1: drafting a performance review report
In this task, participants were instructed to imagine themselves as a team leader at Optima Manufacturing, a company specializing in high-quality electronic components. The objective was to draft a performance review report for their subordinate, Jordan Smith, based on specific performance metrics. Participants were directed

to create a balanced and supportive report that provides insights into Jordan's performance, emphasizing both strengths and areas for growth, with the overall goal of aiding Jordan's professional development. Participants were randomly assigned to one of two conditions: human-ChatGPT collaboration ($N = 402$) or human working-alone ($N = 391$). Building upon the same technique employed in Study 1, participants in the human-ChatGPT collaboration condition were guided to collaborate with ChatGPT and finalize their reports. In contrast, participants in the human-working-alone condition completed this task independently, without any assistance from ChatGPT. Following task completion, both groups answered a series of survey questions that measure their psychological experiences during the task and check the effectiveness of our manipulation.

Task 2: brainstorming for product improvement
Upon completing Task 1, participants from both conditions were presented with a brainstorming task. In this task, they were instructed to think creatively as a product manager and generate innovative ideas to improve an Interactive Whiteboard in three minutes. After this, participants answered another series of questions that asked about their psychological experiences during the brainstorming task and collected their demographic information.

Post-task questionnaires
In Study 2, we employed the same questionnaire utilized in Study 1, maintaining consistent measures for participants' perceived sense of personal control (Cronbach's $\alpha = 0.89$ & 0.90), intrinsic motivation (Cronbach's $\alpha = 0.95$ & 0.95), and feelings of boredom (Cronbach's $\alpha = 0.87$ & 0.89) following Task 1 and Task 2, respectively. Demographic information was similarly gathered at the end of the questionnaire. Due to observed low reliability in the boredom scale during Study 1, we opted for the four-item scale sourced from previous work[52] in Study 2, specifically designed to capture boredom as a dynamic state (e.g., "I found that task boring").

Manipulation check
Similar to Study 1, participants indicated how they completed Task 1, choosing between "I wrote the performance review report purely on my own" and "I collaborated with an artificial intelligence (AI) tool in writing the performance review report". The chi-square test of independence revealed a significant relationship between condition and reported collaboration method (χ^2 (1) = 567.74, $p < .001$). In the human-AI collaboration condition, 92.3% of participants reported collaborating with AI, compared to only 7.7% in the human-working-alone condition. This significant disparity confirms the effective manipulation of the collaboration method.

Assessment of Task 1 performance
We employed the Linguistic Inquiry and Word Count (LIWC) software, a robust tool for text analysis that quantifies psychological and linguistic traits within texts[41]. All language dimensions assessed in the paper were drawn from the standard LIWC2015 dictionary, which allocates each word and word combination into specific linguistic categories.

We determined three specific dimensions that are relevant in evaluating the quality of outputs during the Product Improvement Task according to the task requirement. First, word count was assessed as an indicator of the performance report's thoroughness. Longer reports typically provide a more comprehensive analysis of the subordinate's performance, offering detailed feedback and suggestions for improvement. Second, we assessed analytical content, which reflects the extent to which reports provide constructive feedback. Higher scores in this dimension suggest that the report is logically organized and thoughtfully articulated, which is crucial for creating a balanced and objective performance review. Finally, the prosocial dimension captures the extent to which the language used in the reports is supportive and encouraging, which is essential for a performance review aimed at professional development. Higher scores in this dimension indicate a positive tone that can motivate the subordinate to improve performance.

Assessment of Task 2 performance
To assess the novelty and usefulness of ideas generated, we recruited 15 undergraduate students following the same recruitment procedures of Study 1. Each participant's response was evaluated by a randomly selected subset of 5 out of 15 raters, who approximately evaluate 243 to 294 responses. We provided raters with detailed guidelines to ensure assessment accuracy and uniformity. Their ratings met standard cutoffs for interrater reliability (novelty: ICC(1, 5) = 0.82; usefulness: ICC(1, 5) = 0.80), so we averaged to form a unified measure of each response's novelty and usefulness. For idea quantity, one research assistant who was blind to the study counted the number of ideas that each participant generated, which was then cross verified by a second research assistant.

Study 3

The objective of Study 3 is to further replicate and substantiate the findings from Studies 1 and 2 by using a similar experimental design but introducing new task scenarios. For this study, the tasks include "Composing an Email" for Task 1 and "Idea Generation on Product Promotion" for Task 2. Consistent with the approach in Study 2, we have calculated the required sample size as 788 to detect the smallest effect size of theoretical relevance for our research (α error probability = 0.05, power = 0.80). This adjustment ensures that the study is adequately powered to identify subtle yet significant effects, further strengthening the robustness and generalizability of our findings in human-AI interaction research.

Participants and procedures
Following the same prescreen criteria in Study 2, we collected responses from 793 participants, with none being excluded from the study. Of participants, 49.3% identified as women ($N = 391$). Approximately 73.3% of respondents held an associate degree or higher ($N = 581$). The participants' average age was 38.01 years ($SD = 10.11$). Upon successful survey completion, participants were compensated with 2.25 British pounds.

Task 1: composing an email
In this task, participants were assigned the responsibility of team manager and were required to compose a welcoming email to introduce a new colleague to the team. Building upon the same technique employed in Study 1 and Study 2, participants in the human-ChatGPT condition ($N = 405$) were guided to collaborate with ChatGPT to compose the email. In contrast, participants in the human-working-solo condition ($N = 388$) completed this task independently. Following task completion, both groups answered a series of survey questions that measure their psychological experiences during the task and check the effectiveness of our manipulation.

Task 2: Idea generation on product promotion
Upon completing Task 1, participants from both conditions were presented with a brainstorming task which they worked on independently. In this task, they were instructed to generate numerous innovative marketing promotion ideas for a specified cleaning product, aiming to captivate customers' interest. After this, participants answered another series of questions that asked about their psychological experiences during the brainstorming task and collected their demographic information.

Post-task questionnaires
We employed the same questionnaire utilized in Study 2, maintaining consistent measures for participants' perceived sense of personal control (Cronbach's $\alpha = 0.89$ & 0.89), intrinsic motivation (Cronbach's $\alpha = 0.96$ & 0.94), and feelings of boredom (Cronbach's $\alpha = 0.91$ & 0.89) following Task 1 and Task 2, respectively. Demographic information was similarly gathered at the end of the questionnaire.

Manipulation check
Similar to Study 1 and 2, participants indicated how they completed Task 1, choosing between "I composed the email purely on my own" and "I collaborated with an artificial intelligence (AI) tool in composing the email". The chi-square test of independence revealed a significant relationship between condition and reported collaboration method (χ^2 (1) = 644.75, $p < .001$). In the human-AI collaboration condition, 94.6% of participants reported collaborating with AI, compared to only 5.4% in the human-working-alone condition. Similarly, in the human-working-alone condition, 95.6% reported no collaboration with AI, with only 4.4% indicating any AI involvement. These results confirm that the manipulation of the collaboration method was effectively implemented.

Assessment of Task 1 performance
Using the dictionary-based text analysis tool LIWC, we focused on three specific dimensions: word count, affiliation content, and social orientation. These dimensions are particularly appropriate for evaluating the completeness, team-oriented focus, and interpersonal warmth of the emails. First, word count helps ensure that the email provides sufficient effort to craft comprehensive and engaging introductions. Second, affiliation-oriented dimension reflects the integration of a new team member positively. High scores in this dimension suggest that an email effectively fosters a sense of unity and highlights the importance of team spirit among team members. Finally, the social dimension indicates the extent to which an email is warm, welcoming and personable, which conveys friendliness and sets a positive tone for the new employee.

Assessment of Task 2 performance
Following the same recruiting and evaluating procedures of Study 2, the novelty and usefulness of ideas generated by each participant were scored by 5 independent raters who were randomly selected from a pool of 15. The ratings met standard cutoffs for interrater reliability (novelty: ICC(1, 5) = 0.80; usefulness: ICC(1, 5) = 0.69), so we averaged to form a unified measure of each response's novelty and usefulness. For idea quantity, one research assistant who was blind to the study counted the number of ideas that each participant generated, which was then cross verified by a second research assistant.

Study 4

Building on Studies 1 to 3, Study 4 employed a two-task design to replicate prior findings and explore two additional conditions: Transitioning from solo work to GenAI collaboration and completing consecutive tasks with GenAI collaboration. To control for potential confounding effects, such as task type, time constraints, and task order, we made several key modifications to the study design. First, unlike Studies 1 to 3, where we used two different tasks (one text-generation task and one brainstorming task), Study 4 included two comparable text-generation tasks to ensure consistency. Specifically, we used two tasks from Study 1 and Study 3—"Composing a Facebook Post" and "Writing a Welcome Email". These tasks were chosen because of their similar cognitive demands, as indicated by participants' cognitive load ratings in previous studies. Both tasks involve content generation, ensuring a fundamental similarity in their nature. Second, participants in Study 4 were not subject to time limits, allowing them to complete the tasks at their own pace. Third, to address potential order effects, we counterbalanced the presentation of the two tasks by randomizing the task order within the survey flow.

For each task, participants had an equal probability (50%) of being assigned to either complete the task with the assistance of ChatGPT or complete it independently. This design resulted in a 2×2 mixed factorial structure,

with one between-subjects factor (Collaborating with ChatGPT vs. Solo) and one within-subjects factor (Task 1 vs. Task 2), yielding four experimental conditions: Collab-Collab, Collab-Solo, Solo-Solo, and Solo-Collab. Additionally, to mitigate the potential influence of task sequence, we randomly assigned two tasks to participants. Specifically, for each of the four conditions, half of the participants completed "Writing a Welcome Email" as Task 1 and "Composing a Facebook Post' as Task 2, while the other half completed the tasks in the reverse order ("Composing a Facebook Post" as Task 1 and "Writing a Welcome Email" as Task 2).

After completing each task, participants reported their sense of control, intrinsic motivation, and feelings of boredom on 7-point scales. This setup allowed us to examine variations in these psychological experiences across tasks and conditions. The analyses included mixed-design repeated measures ANOVAs to test interaction effects and paired-samples *t*-tests to assess within-individual changes.

Participants and procedures

The study was conducted via the platform Prolific, targeting participants who are UK residents, have English as their first language, are employed in either full-time or part-time positions, and are aged between 18 and 60 years. Participants who participated in our previous studies were excluded. The study takes approximately 10 min, with each eligible participant receiving compensation of £1.50 upon successful completion.

As with previous studies, we determined that "$d = 0.20$" represents the smallest effect size of theoretical interest for our study. Using this effect size, a desired statistical power of 80%, and a significance level of $\alpha = 0.05$, we calculated a required sample size of 393 participants per group, resulting in a total sample size of 1,572 participants. This sample size ensures sufficient power to detect small but meaningful effects, balancing the risks of Type I and Type II errors while adhering to rigorous methodological standards.

Data collection was initially halted at 1,572 submissions, which corresponded to the predetermined sample size based on our priori power analysis. Upon reviewing the data, we identified and excluded 415 submissions (approximately 26%) that failed the integrity check, as these participants reported using artificial intelligence when they were instructed to work independently. To bolster our statistical power, we decided to extend data collection to reach beyond the initial sample size, ultimately collecting an additional 600 submissions, of which 133 failed the integrity check (22%). Consequently, the total number of submissions reached 2,172. After rigorously applying the exclusion criteria to remove those who did not adhere to the study's instructions (548, approximately 25%), we were left with a final sample size of 1,624 participants. This refined dataset forms the basis of our analysis.

Post-task questionnaires

We employed the same questionnaire utilized in Study 2 and 3, maintaining consistent measures for participants' perceived sense of personal control (Cronbach's $\alpha = 0.91$ & 0.90), intrinsic motivation (Cronbach's $\alpha = 0.96$ & 0.96), and feelings of boredom (Cronbach's $\alpha = 0.91$ & 0.90) following Task 1 and Task 2, respectively. Demographic information was gathered at the end of the questionnaire.

Manipulation check

Participants indicated how they completed Task 1 by selecting either "I composed the email purely on my own" or "I collaborated with an artificial intelligence (AI) tool in composing the email". To evaluate the effectiveness of the manipulation, we focused on cases from the Collab-Solo and Solo-Solo conditions. A chi-square test of independence revealed a significant relationship between condition and reported collaboration method ($\chi^2(1) = 687.97$, $p < .001$). In the Solo-Solo condition, 99.8% of participants reported no collaboration with AI, with only 0.2% indicating AI involvement. Conversely, in the Collab-Solo condition, 93.9% of participants reported collaborating with GenAI, while 6.1% indicated working independently. These results confirm that the collaboration method was successfully manipulated as intended.

Assessment of task performance

Task performance was evaluated using the dictionary-based text analysis tool LIWC, focusing on three key dimensions: word count, analytical content, and positive tone. Word count served as a measure of the detail and completeness of the text. Analytical content assessed the logical structure and clarity of the writing, essential for effectively conveying information and achieving the task's purpose. Positive tone evaluated the degree of engaging, professional, and persuasive communication, particularly relevant for tasks like crafting a welcoming email or a promotional social media post.

Preregistrations

All four studies were preregistered at the website https://aspredicted.org. The preregistrations are now publicly available and can be accessed at https://aspredicted.org/2c56-fvkk.pdf (Study 1), https://aspredicted.org/yi6vg.pdf (Study 2), https://aspredicted.org/rmp5-28t9.pdf (Study 3), and https://aspredicted.org/crnf-qw7d.pdf (Study 4) respectively. We also report any deviations from our preregistrations in Table 1 in our Supplementary Materials.

Data availability

Data, study materials, and data analyses codes for the study are publicly available in the Open Science Framework (OSF). All study materials, anonymized raw data, raters' evaluation records, cleaned datasets used for analyses (in .sav and .csv formats), and SPSS Syntax files for data analyses can be accessed at https://osf.io/xyjwh.

Received: 30 August 2024; Accepted: 11 April 2025

Published online: 29 April 2025

References

1. Raisch, S. & Krakowski, S. Artificial intelligence and management: the automation–augmentation paradox. *Acad. Manag Rev.* **46**, 192–210 (2021).
2. Wilson, H. J. & Daugherty, P. R. Collaborative intelligence: humans and AI are joining forces. *Harv. Bus. Rev.* **96**, 114–123 (2018).
3. Brynjolfsson, E., Li, D. & Raymond, L. Generative AI at work. *Q. J. Econ.* **140**, 889–942 (2025).
4. Noy, S. & Zhang, W. Experimental evidence on the productivity effects of generative artificial intelligence. *Science* **381**, 187–192 (2023).
5. Microsoft *AI is already changing work—Microsoft included.* Retrieved January 23, from (2025). https://www.microsoft.com/en-us/worklab/ai-is-already-changing-work-microsoft-included
6. Microsoft *AI at work is here—Now comes the hard part.* Retrieved January 23, from (2025). https://www.microsoft.com/en-us/worklab/work-trend-index/ai-at-work-is-here-now-comes-the-hard-part
7. Gagné, M. et al. Understanding and shaping the future of work with self-determination theory. *Nat. Rev. Psychol.* **1**, 378–392 (2022).
8. Raisch, S. & Fomina, K. Combining human and artificial intelligence: hybrid problem-solving in organizations. *Acad. Manag Rev.* https://doi.org/10.5465/amr.2021.0421 (2024).
9. Parker, S. K. & Grote, G. Automation, algorithms, and beyond: why work design matters more than ever in a digital world. *Appl. Psychol.* **71**, 1171–1204 (2022).
10. Lee, B. C. & Chung, J. An empirical investigation of the impact of ChatGPT on creativity. *Nat. Hum. Behav.* **8**, 1906–1914 (2024).
11. Doshi, A. R. & Hauser, O. P. Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Sci. Adv.* **10**, eadn5290 (2024).
12. Dell'Acqua, F. et al. Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality. *Harvard Business School Technology & Operations Mgt. Unit Working Paper No. 24–013; The Wharton School Research Paper.* Available https://ssrn.com/abstract=4573321 (2023).
13. Doshi, A. R., Bell, J. J., Mirzayev, E. & Vanneste, B. S. Generative artificial intelligence and evaluating strategic decisions. *Strateg Manag J.* **46**, 583–610 (2025).
14. Gilardi, F., Alizadeh, M. & Kubli, M. ChatGPT outperforms crowd workers for text-annotation tasks, *Proc. Natl. Acad. Sci. U.S.A.* 120, e2305016120 (2023).
15. Peng, S., Kalliamvakou, E., Cihon, P. & Demirer, M. The impact of AI on developer productivity: Evidence from github copilot. *arXiv preprint at* (2023). https://arxiv.org/abs/2302.06590*v1*.
16. Sharma, A., Lin, I. W., Miner, A. S., Atkins, D. C. & Althoff, T. Human–AI collaboration enables more empathic conversations in text-based peer-to-peer mental health support. *Nat. Mach. Intell.* **5**, 46–57 (2023).
17. Jia, N., Luo, X., Fang, Z. & Liao, C. When and how artificial intelligence augments employee creativity. *Acad. Manag J.* **67**, 5–32 (2024).
18. Deci, E. L. & Ryan, R. M. *Intrinsic Motivation and self-determination in Human Behavior* (New York Plenum, 1985).
19. Ryan, R. M. & Deci, E. L. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* **55**, 68–78 (2000).
20. Deci, E. L., Olafsen, A. H. & Ryan, R. M. Self-determination theory in work organizations: the state of a science. *Annu. Rev. Organ. Psychol. Organ. Behav.* **4**, 19–43 (2017).
21. Amabile, T. M. A model of creativity and innovation in organizations. *Res. Organ. Behav.* **10**, 123–167 (1988).
22. Oldham, G. R. & Cummings, A. Employee creativity: personal and contextual factors at work. *Acad. Manag J.* **39**, 607–634 (1996).
23. Zhou, J. & Hoever, I. J. Research on workplace creativity: A review and Redirection. *Annu. Rev. Organ. Psychol. Organ. Behav.* **1**, 333–359 (2014).
24. Lu, J. G., Akinola, M. & Mason, M. F. Switching on creativity: task switching can increase creativity by reducing cognitive fixation. *Organ. Behav. Hum. Decis. Process.* **139**, 63–75 (2017).
25. Li, J., Huang, C. C., Wu, M. C. & Huang, Y. H. The impact of an intelligent writing assistant on writing proficiency and self-efficacy. *J. Educ. Technol. Soc.* **23**, 82–95 (2020).
26. Sun, M. & Wang, Y. The effectiveness of an AI-powered writing coach in improving writing quality and self-efficacy. *Int. J. Hum. Comput Stud.* **138**, 102394 (2020).
27. Zuckerman, M., Porac, J., Lathin, D. & Deci, E. L. On the importance of self-determination for intrinsically-motivated behavior. *Pers. Soc. Psychol. Bull.* **4**, 443–446 (1978).
28. Ryan, R. M. & Deci, E. L. Intrinsic and extrinsic motivations: classic definitions and new directions. *Contemp. Educ. Psychol.* **25**, 54–67 (2000).
29. Dell'Acqua, F. Falling asleep at the wheel: Human/AI collaboration in a field experiment on HR recruiters [Working Paper]. New York, NY: Columbia Business School. (2021).
30. Gallagher, S. Philosophical conceptions of the self: implications for cognitive science. *Trends Cogn. Sci.* **4**, 14–21 (2000).
31. Faas, C., Bergs, R., Sterz, S., Langer, M. & Feit, A. M. Give me a choice: The consequences of restricting choices through AI-support for perceived autonomy, motivational variables, and decision performance. *arXiv preprint at* (2024). https://arxiv.org/abs/2410.07728
32. Lu, W. Inevitable challenges of autonomy: ethical concerns in personalized algorithmic decision-making. *Humanit. Soc. Sci. Commun.* **11**, 1–9 (2024).
33. Nguyen, Q. N., Sidorova, A. & Torres, R. User interactions with chatbot interfaces vs. Menu-based interfaces: an empirical study. *Comput. Hum. Behav.* **128**, 107093 (2022).
34. Shin, J. & Grant, A. M. Bored by interest: how intrinsic motivation in one task can reduce performance on other tasks. *Acad. Manag J.* **62**, 415–436 (2019).
35. Fisher, C. D. Boredom at work: A neglected concept. *Hum. Relat.* **46**, 395–417 (1993).
36. Mael, F. & Jex, S. Workplace boredom: an integrative model of traditional and contemporary approaches. *Group. Organ. Manag.* **40**, 131–159 (2015).
37. van Hooff, M. L. & van Hooft, E. A. Boredom at work: proximal and distal consequences of affective work-related boredom. *J. Occup. Health Psychol.* **19**, 348 (2014).
38. van Hooff, M. L. & van Hooft, E. A. Boredom at work: towards a dynamic spillover model of need satisfaction, work motivation, and work-related boredom. *Eur. J. Work Organ. Psychol.* **26**, 133–148 (2017).
39. Van Tilburg, W. A. & Igou, E. R. On boredom: lack of challenge and meaning as distinct boredom experiences. *Motiv Emot.* **36**, 181–194 (2012).
40. Guilford, J. P. *The Nature of Human Intelligence* (McGraw-Hill, 1967).
41. Pennebaker, J. W., Booth, R. J., Boyd, R. L. & Francis, M. E. *LIWC 2015 Operator's Manual* (Pennebaker Conglomerates Inc, 2015).
42. Berg, J., Raj, M. & Seamans, R. Capturing value from artificial intelligence. *Acad. Manag Discov.* **9**, 424–428 (2023).
43. Feyer, F. K. Creating a model of cross-task motivation—A meta-narrative review of the literature on dynamic motivation. *Front. Psychol.* **14**, 1193186 (2023).

44. Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods*. **39**, 175–191 (2007).
45. Faul, F., Erdfelder, E., Buchner, A. & Lang, A. G. Statistical power analyses using G* power 3.1: tests for correlation and regression analyses. *Behav. Res. Methods*. **41**, 1149–1160 (2009).
46. Velez, Y. R. Leveraging large Language models for public opinion and political psychology research: opportunities and challenges. *OSF Preprints*. https://doi.org/10.17605/osf.io/5sxdt (2024).
47. Greenaway, K. H. et al. From we to me: group identification enhances perceived personal control with consequences for health and well-being. *Pers. Soc. Psychol.* **109**, 53–74 (2015).
48. Koo, T. K. & Li, M. Y. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* **15**, 155–163 (2016).
49. Shrout, P. E. & Fleiss, J. L. Intraclass correlations: uses in assessing rater reliability. *Psychol. Bull.* **86**, 420–428 (1979).
50. Simonsohn, U. Small telescopes: detectability and the evaluation of replication results. *Psychol. Sci.* **26**, 559–569 (2015).
51. Lakens, D. Equivalence tests: A practical primer for t tests, correlations, and meta-analyses. *Soc. Psychol. Pers. Sci.* **8**, 355–362 (2017).
52. van Hooft, E. A. & van Hooff, M. L. The state of boredom: frustrating or depressing? *Motiv Emot.* **42**, 931–946 (2018).

Author contributions

S.W. and Y.L. initiated the project, designed all experiments, and wrote the main manuscript. M.R. and S.C. prepared the study materials (Studies 1-3), coded all textual data, and contributed to writing the methods section (Studies 1-3). X.X. provided feedback on the design and execution of the studies and secured funding to support this research. All authors reviewed and approved the final manuscript.

Funding

This research was supported by multiple research projects funded by the National Natural Science Foundation of China (Grant Numbers: 72201241, 72302214, and 72232009), the Humanities and Social Science Research Project funded by the Ministry of Education of China (Grant Number: 22YJC630094), and the Zhejiang Provincial Philosophy and Social Sciences Planning Project (Grant Numbers: 23YJRC03ZD-2YB and 23YJRC03ZD-3YB).

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-98385-2.

Correspondence and requests for materials should be addressed to Y.L.

Reprints and permissions information is available at www.nature.com/reprints.

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

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2025