How and For Whom Using Generative AI Affects Creativity: A Field Experiment

Supplemental Materials

- Tables S1–S4 summarize the additional analyses using each item from the cognitive job resources scale. A summary of these analyses is provided below:
  - As shown in Table S1, metacognitive strategies significantly moderated the effect of LLM assistance on the overall scale of cognitive job resources ($\gamma = 0.62$, $SE = 0.16$, $p < .001$) and each individual item ($\gamma = 0.50$, $SE = 0.19$, $p = .008$; $\gamma = 0.58$, $SE = 0.19$, $p = .003$; $\gamma = 0.79$, $SE = 0.18$, $p < .001$; $\gamma = 0.62$, $SE = 0.19$, $p = .002$).
  - Table S2 demonstrates that the overall scale of cognitive job resources ($\gamma = 0.23$, $SE = 0.06$, $p < .001$) and each item were significantly associated with supervisor-rated creativity ($\gamma = 0.15$, $SE = 0.05$, $p = .007$; $\gamma = 0.15$, $SE = 0.05$, $p = .005$; $\gamma = 0.20$, $SE = 0.06$, $p < .001$; $\gamma = 0.17$, $SE = 0.05$, $p = .002$).
  - Table S3 examines the effects of cognitive job resources on external raters' novelty ratings. Except for item 2 (i.e., opportunities for task-switching, $\gamma = 0.02$, $SE = 0.07$, $p = .798$), the overall scale ($\gamma = 0.16$, $SE = 0.08$, $p = .04$) and the remaining items were significantly or marginally related to novelty rated by external raters ($\gamma = 0.17$, $SE = 0.07$, $p = .011$; $\gamma = 0.17$, $SE = 0.07$, $p = .014$; $\gamma = 0.12$, $SE = 0.07$, $p = .074$).
  - Table S4 focuses on the effects of cognitive job resources on external raters' usefulness ratings. Except for item 2 (i.e., opportunities for task-switching, $\gamma = 0.09$, $SE = 0.06$, $p = .113$), the overall scale ($\gamma = 0.22$, $SE = 0.07$, $p = .001$) and the remaining items were significantly related to usefulness rated by external raters ($\gamma = 0.19$, $SE = 0.06$, $p = .001$; $\gamma = 0.23$, $SE = 0.06$, $p < .001$; $\gamma = 0.13$, $SE = 0.06$, $p = .022$). These analyses yielded results similar to those from the overall scale score of cognitive job resources, suggesting that our observed effects are robust across different components of the construct.
- Table S5 presents hypothesis testing results without control variables, which remain consistent with those reported in the main text with control variables.
- S6 summarizes robustness tests for the moderation effects of metacognitive strategies.
- S7 evaluates the practical significance of the conditional indirect effect of LLM assistance on employee creativity via cognitive job resources.
- S8 compares model fit indices with or without metacognitive strategies as a moderator.

Table S1: The Moderating Effect of Metacognitive Strategies on the Relationship Between LLM Assistance and Cognitive Job Resources (Overall Scale and Individual Scale Items)

	Cognitive Job Resources (Overall Scale)	Cognitive Job Resources (Item 1)	Cognitive Job Resources (Item 2)	Cognitive Job Resources (Item 3)	Cognitive Job Resources (Item 4)
Intercept	3.43***(0.07)	3.38***(0.09)	3.41***(0.09)	3.47***(0.08)	3.45***(0.08)
LLM assistance condition	0.63***(0.10)	0.74***(0.12)	0.56***(0.12)	0.65***(0.12)	0.60***(0.12)
Metacognitive strategies	0.05(0.15)	0.10(0.18)	0.11(0.18)	-0.07(0.17)	0.08(0.18)
LLM assistance condition × Metacognitive strategies	**0.62***(0.16)**	**0.50**(0.19)**	**0.58**(0.19)**	**0.79***(0.18)**	**0.62**(0.19)**
Heuristic task characteristics: unclear means	0.11(0.11)	0.20(0.12)	0.04(0.13)	0.14(0.12)	0.04(0.13)
Heuristic task characteristics: unclear ends	0.03(0.10)	-0.03(0.11)	0.11(0.12)	-0.00(0.11)	0.05(0.12)
Past job performance	0.01(0.01)	0.01(0.02)	0.02(0.02)	0.02(0.02)	0.01(0.02)
Job insecurity	0.01(0.06)	-0.01(0.06)	0.02(0.07)	0.04(0.06)	-0.00(0.07)
Creative self-efficacy	0.11(0.06)	0.12(0.07)	0.16*(0.07)	0.08(0.07)	0.09(0.07)
Intrinsic motivation	0.12(0.10)	0.10(0.12)	0.10(0.12)	0.17(0.12)	0.09(0.12)
Variance (Level 3 intercept)	0.00(0.00)	0.00(0.00)	0.00(0.00)	0.00(0.00)	0.00(0.00)
Variance (Level 2 intercept)	0.01(0.02)	0.02(0.03)	0.02(0.03)	0.03(0.03)	0.00(0.00)
Variance (Level 1 residual)	0.55(0.05)	0.74(0.07)	0.78(0.10)	0.68(0.06)	0.79(0.07)
Pseudo R^2	0.06	0.03	0.04	0.06	0.04

Note. $N = 250$. * $p < .05$; ** $p < .01$. *** $p < .001$. Standard errors are in parentheses. LLM Assistance Condition: 0 = Control (no LLM assistance); 1 = Experimental (with LLM assistance). All continuous variables are mean-centered. Values in bold are relevant to hypothesis testing. *Item 1:* I have the opportunity to take a mental break when tasks require a lot of concentration. *Item 2:* I have the opportunity to vary complex tasks with simple tasks. *Item 3:* I have access to useful information to help solve complex tasks. *Item 4:* I have the opportunity to determine my own work method.

Table S2: The Effect of Cognitive Job Resources (Overall Scale and Individual Scale Items) on Supervisor-Rated Creativity

	Creativity (supervisor rating)	Creativity (supervisor rating)	Creativity (supervisor rating)	Creativity (supervisor rating)	Creativity (supervisor rating)
Intercept	2.40***(0.23)	2.67***(0.20)	2.66***(0.20)	2.48***(0.21)	2.60***(0.20)
LLM assistance condition	0.70***(0.11)	0.74***(0.11)	0.76***(0.11)	0.72***(0.11)	0.74***(0.11)
Cognitive job resources (overall scale)	**0.23***(0.06)**				
Cognitive job resources (Item 1)		**0.15**(0.05)**			
Cognitive job resources (Item 2)			**0.15**(0.05)**		
Cognitive job resources (Item 3)				**0.20***(0.06)**	
Cognitive job resources (Item 4)					**0.17**(0.05)**
Metacognitive strategies	0.01(0.15)	0.01(0.15)	0.01(0.15)	0.04(0.15)	0.01(0.15)
LLM assistance condition × Metacognitive strategies	-0.17(0.17)	-0.10(0.17)	-0.11(0.17)	-0.19(0.17)	-0.13(0.17)
Heuristic task characteristics: unclear means	-0.03(0.10)	-0.04(0.11)	-0.01(0.11)	-0.04(0.11)	-0.01(0.11)
Heuristic task characteristics: unclear ends	-0.02(0.10)	-0.01(0.10)	-0.03(0.10)	-0.01(0.10)	-0.02(0.10)
Past job performance	0.01(0.01)	0.01(0.02)	0.01(0.02)	0.01(0.01)	0.01(0.02)
Job insecurity	0.01(0.06)	0.01(0.06)	0.01(0.06)	-0.00(0.06)	0.01(0.06)
Creative self-efficacy	0.47***(0.06)	0.48***(0.06)	0.48***(0.06)	0.48***(0.06)	0.49***(0.06)
Intrinsic motivation	-0.03(0.10)	-0.02(0.10)	-0.02(0.10)	-0.04(0.10)	-0.02(0.10)
Variance (Level 3 intercept)	0.00(0.00)	0.00(0.00)	0.00(0.00)	0.00(0.00)	0.00(0.00)
Variance (Level 2 intercept)	0.03(0.04)	0.04(0.03)	0.03(0.03)	0.02(0.03)	0.03(0.03)
Variance (Level 1 residual)	0.53(0.05)	0.54(0.05)	0.54(0.05)	0.54(0.05)	0.54(0.05)
Pseudo R^2	0.05	0.03	0.03	0.05	0.04

Note. $N = 250$. * $p < .05$; ** $p < .01$. *** $p < .001$. Standard errors are in parentheses. LLM Assistance Condition: 0 = Control (no LLM assistance); 1 = Experimental (with LLM assistance). All continuous variables are mean-centered. Values in bold are relevant to hypothesis testing. *Item 1:* I have the opportunity to take a mental break when tasks require a lot of concentration. *Item 2:* I have the opportunity to vary complex tasks with simple tasks. *Item 3:* I have access to useful information to help solve complex tasks. *Item 4:* I have the opportunity to determine my own work method.

Table S3: The Effect of Cognitive Job Resources (Overall Scale and Individual Scale Items) on External Raters' Ratings of Novelty

	Novelty (external raters' ratings)	Novelty (external raters' ratings)	Novelty (external raters' ratings)	Novelty (external raters' ratings)	Novelty (external raters' ratings)
Intercept	2.27***(0.28)	2.24***(0.24)	2.76***(0.24)	2.23***(0.25)	2.41***(0.24)
LLM assistance condition	0.13(0.13)	0.11(0.13)	0.23(0.13)	0.13(0.13)	0.16(0.13)
Cognitive job resources (overall scale)	**0.16*(0.08)**				
Cognitive job resources (Item 1)		**0.17*(0.07)**			
Cognitive job resources (Item 2)			**0.02(0.07)**		
Cognitive job resources (Item 3)				**0.17*(0.07)**	
Cognitive job resources (Item 4)					**0.12(0.07)**
Metacognitive strategies	-0.51**(0.19)	-0.51**(0.18)	-0.50**(0.19)	-0.48**(0.18)	-0.51**(0.19)
LLM assistance condition × Metacognitive strategies	0.45*(0.21)	0.46*(0.20)	0.54**(0.21)	0.42*(0.21)	0.48*(0.21)
Heuristic task characteristics: unclear means	-0.02(0.13)	-0.04(0.13)	-0.00(0.13)	-0.03(0.13)	-0.01(0.13)
Heuristic task characteristics: unclear ends	0.00(0.12)	0.01(0.12)	0.01(0.12)	0.01(0.12)	0.00(0.12)
Past job performance	0.03(0.02)	0.03(0.02)	0.03(0.02)	0.03(0.02)	0.03(0.02)
Job insecurity	-0.00(0.07)	0.00(0.07)	-0.00(0.07)	-0.01(0.07)	-0.00(0.07)
Creative self-efficacy	0.13(0.08)	0.12(0.08)	0.14(0.08)	0.13(0.08)	0.13(0.08)
Intrinsic motivation	0.20(0.13)	0.20(0.13)	0.22(0.13)	0.19(0.13)	0.21(0.13)
Variance (Level 3 intercept)	0.00(0.00)	0.00(0.00)	0.00(0.00)	0.00(0.00)	0.00(0.00)
Variance (Level 2 intercept)	0.00(0.00)	0.00(0.02)	0.00(0.00)	0.00(0.00)	0.00(0.00)
Variance (Level 1 residual)	0.84(0.08)	0.83(0.08)	0.84(0.08)	0.83(0.07)	0.84(0.08)
Pseudo R^2	0.02	0.03	0.00	0.02	0.01

Note. $N = 250$. * $p < .05$; ** $p < .01$. *** $p < .001$. Standard errors are in parentheses. LLM Assistance Condition: 0 = Control (no LLM assistance); 1 = Experimental (with LLM assistance). All continuous variables are mean-centered. Values in bold are relevant to hypothesis testing. *Item 1:* I have the opportunity to take a mental break when tasks require a lot of concentration. *Item 2:* I have the opportunity to vary complex tasks with simple tasks. *Item 3:* I have access to useful information to help solve complex tasks. *Item 4:* I have the opportunity to determine my own work method.

Table S4: The Effect of Cognitive Job Resources (Overall Scale and Individual Scale Items) on External Raters' Ratings of Usefulness

	Usefulness (external raters' ratings)	Usefulness (external raters' ratings)	Usefulness (external raters' ratings)	Usefulness (external raters' ratings)	Usefulness (external raters' ratings)
Intercept	2.17***(0.24)	2.28***(0.21)	2.61***(0.21)	2.11***(0.22)	2.47***(0.21)
LLM assistance condition	0.02(0.12)	0.01(0.12)	0.10(0.11)	0.01(0.11)	0.08(0.11)
Cognitive job resources (overall scale)	**0.22**(0.07)**				
Cognitive job resources (Item 1)		**0.19**(0.06)**			
Cognitive job resources (Item 2)			**0.09(0.06)**		
Cognitive job resources (Item 3)				**0.23***(0.06)**	
Cognitive job resources (Item 4)					**0.13*(0.06)**
Metacognitive strategies	-0.43**(0.16)	-0.44**(0.16)	-0.43**(0.16)	-0.40*(0.16)	-0.43**(0.16)
LLM assistance condition × Metacognitive strategies	0.32(0.18)	0.36*(0.18)	0.40*(0.18)	0.27(0.18)	0.37*(0.18)
Heuristic task characteristics: unclear means	0.14(0.11)	0.12(0.11)	0.16(0.12)	0.13(0.11)	0.16(0.12)
Heuristic task characteristics: unclear ends	-0.18(0.11)	-0.16(0.11)	-0.18(0.11)	-0.17(0.10)	-0.18(0.11)
Past job performance	0.02(0.02)	0.02(0.02)	0.02(0.02)	0.02(0.02)	0.02(0.02)
Job insecurity	0.02(0.06)	0.03(0.06)	0.03(0.06)	0.02(0.06)	0.03(0.06)
Creative self-efficacy	0.15*(0.07)	0.15*(0.07)	0.16*(0.07)	0.16*(0.07)	0.16*(0.07)
Intrinsic motivation	0.15(0.11)	0.15(0.11)	0.16(0.11)	0.13(0.11)	0.16(0.11)
Variance (Level 3 intercept)	0.00(0.00)	0.00(0.00)	0.00(0.00)	0.00(0.00)	0.00(0.00)
Variance (Level 2 intercept)	0.00(0.00)	0.00(0.00)	0.00(0.00)	0.00(0.00)	0.00(0.00)
Variance (Level 1 residual)	0.64(0.06)	0.64(0.06)	0.66(0.06)	0.63(0.07)	0.66(0.06)
Pseudo R^2	0.04	0.04	0.01	0.06	0.02

Note. $N = 250$. * $p < .05$; ** $p < .01$. *** $p < .001$. Standard errors are in parentheses. LLM Assistance Condition: 0 = Control (no LLM assistance); 1 = Experimental (with LLM assistance). All continuous variables are mean-centered. Values in bold are relevant to hypothesis testing. *Item 1:* I have the opportunity to take a mental break when tasks require a lot of concentration. *Item 2:* I have the opportunity to vary complex tasks with simple tasks. *Item 3:* I have access to useful information to help solve complex tasks. *Item 4:* I have the opportunity to determine my own work method.

Table S5. Results for Hypothesis Testing Without Control Variables

	Model 1: Cognitive Job Resources	Model 2: Creativity (supervisor rating)	Model 3: Novelty (external raters' ratings)	Model 4: Usefulness (external raters' ratings)	Model 5: Creativity (supervisor rating)	Model 6: Novelty (external raters' ratings)	Model 7: Usefulness (external raters' ratings)	Model 8: Cognitive Job Resources	Model 9: Creativity (supervisor rating)	Model 10: Novelty (external raters' ratings)	Model 11: Usefulness (external raters' ratings)
Intercept	3.38***(0.08)	3.06***(0.09)	2.78***(0.09)	2.87***(0.08)	2.15***(0.23)	2.13***(0.26)	2.10***(0.23)	3.39***(0.07)	2.02***(0.25)	2.09***(0.27)	2.00***(0.24)
LLM assistance condition	**0.77***(0.11)**	**1.05***(0.11)**	**0.33**(0.12)**	**0.28**(0.11)**	0.85***(0.12)	0.19(0.13)	0.10(0.12)	0.72***(0.10)	0.84***(0.12)	0.19(0.13)	0.10(0.11)
Cognitive job resources					**0.27**(0.06)**	**0.19**(0.07)**	**0.23***(0.06)**	---	**0.31***(0.07)**	**0.20**(0.08)**	**0.25***(0.07)**
Metacognitive strategies								0.22*(0.11)	0.06(0.12)	-0.26*(0.13)	-0.26*(0.11)
LLM assistance condition × Metacognitive strategies								**0.52**(0.17)**	-0.32(0.18)	0.33(0.20)	0.22(0.18)
Variance (Level 3 intercept)	0.00(0.00)	0.00(0.01)	0.00(0.00)	0.00(0.00)	0.00(0.01)	0.00(0.00)	0.00(0.00)	0.00(0.00)	0.00(0.00)	0.00(0.00)	0.00(0.00)
Variance (Level 2 intercept)	0.02(0.03)	0.05(0.04)	0.00(0.03)	0.00(0.00)	0.04(0.04)	0.01(0.02)	0.00(0.00)	0.01(0.02)	0.04(0.03)	0.00(0.00)	0.00(0.00)
Variance (Level 1 residual)	0.67(0.07)	0.71(0.07)	0.90(0.10)	0.72(0.06)	0.67(0.06)	0.88(0.08)	0.69(0.06)	0.58(0.08)	0.66(0.06)	0.87(0.08)	0.67(0.06)
Pseudo R^2	0.18	0.28	0.03	0.03	0.07	0.03	0.05	0.04	0.08	0.03	0.05

Note. $N = 250$. * $p < .05$; ** $p < .01$. *** $p < .001$. Standard errors are in parentheses. LLM Assistance Condition: 0 = Control (no LLM assistance); 1 = Experimental (with LLM assistance). All continuous variables are mean-centered. Values in bold are relevant to hypothesis testing.

S6. Summary of Robustness Tests for the Moderation Effects of Metacognitive Strategies

We conducted several supplementary robustness tests to assess the robustness of the moderating effects of metacognitive strategies. First, we adopted Nahhas's (2024) approach to run sensitivity analyses. Specifically, we dichotomized the continuous metacognitive strategies variable into binary categories (low vs. high) and performed subgroup analyses. Two types of subgroups were created: one using a mean split and the other using a median split of metacognitive strategies. The results showed that the moderating effects of metacognitive strategies are robust across both dichotomized variables: $\gamma = 0.50$, $SE = 0.20$, $p = .007$ (mean split) and $\gamma = 0.50$, $SE = 0.20$, $p < .001$ (median split). That is, these results are consistent with those obtained using the continuous metacognitive strategies variable.

Second, we examined the moderated effects of metacognitive strategies both with (Model 8 in Table 3) and without control variables (Model 8 in Table S5), and the results were robust regardless of their inclusion. Additionally, considering the positive correlation between metacognitive strategies and intrinsic motivation and the negative correlation between metacognitive strategies and job insecurity, we conducted supplemental analyses to assess the robustness of the moderating effects of metacognitive strategies by including intrinsic motivation and job insecurity as competing moderators. The results showed that the moderating effect of metacognitive strategies remained significant ($\gamma = 0.58$, $SE = 0.25$, $p = .023$), whereas the moderating effect of intrinsic motivation was not significant ($\gamma = 0.04$, $SE = 0.21$, $p = .829$). Similarly, the moderating effect of metacognitive strategies remained significant ($\gamma = 0.49$, $SE = 0.18$, $p = .006$), while the moderating effect of job insecurity was not significant ($\gamma = -0.20$, $SE = 0.11$, $p = .068$).

Third, to further ascertain the moderating effect of metacognitive strategies, we followed the procedure outlined by Bliese and Wang (2020, p. 1277) to calculate the cumulative probability of finding significance given the characteristics of our sample and model. Using their prescribed formula, implemented with the *pnorm* function in R, we found a 97% cumulative probability of observing a significant moderating effect of metacognitive strategies. This high probability reinforces the robustness of our results. As Bliese and Wang (2020, p. 1275) noted, this information "can help promote robust and reliable research by conveying appropriate levels of uncertainty."

S7. The Practical Significance of the Conditional Indirect Effect of LLM Assistance on Employee Creativity via Cognitive Job Resources

To assess practical significance, we followed MacKinnon (2008) and standardized the conditional indirect effect using MacKinnon's formula: ab/σ_Y—the ratio of the indirect effect (ab) to the standard deviation (σ_Y) of the outcome variable (Y). This allows us to interpret the conditional indirect effect as the expected change (in standard deviations) in creativity resulting from LLM assistance, via cognitive job resources. To contextualize our findings, we compared them with the widely cited *online experiment* by Noy and Zhang (2023), which examined the effect of ChatGPT assistance on participants' creative writing performance. They showed a 0.45 *SD* increase in writing performance from LLM assistance.

In our field experiment, we found that when metacognitive strategies are high (+1 *SD*),

- LLM assistance increased supervisor-rated creativity by 0.22 *SD* (95% CI = [0.096, 0.364]) via cognitive job resources.
- LLM assistance increased external raters' ratings of idea novelty by 0.17 *SD* (95% CI = [0.008, 0.339]) via cognitive job resources.
- LLM assistance increased external raters' ratings of idea usefulness by 0.26 *SD* (95% CI = [0.096, 0.436]) via cognitive job resources.

While these effects are smaller than those found by Noy and Zhang (2023), it is important to note that our study was conducted in the field with employees performing complex, real-world tasks—far beyond the scope of their relatively simpler writing tasks. Additionally, when metacognitive strategies increase further (e.g., +1.5 or +2 *SD*), the conditional indirect effects also increase.

S8. Model Fit Indices with or without Metacognitive Strategies as a Moderator

	Cognitive Job Resources (Overall Scale)		Cognitive Job Resources (Item 1)		Cognitive Job Resources (Item 2)		Cognitive Job Resources (Item 3)		Cognitive Job Resources (Item 4)	
	AIC	*BIC*	*AIC*	*BIC*	*AIC*	*BIC*	*AIC*	*BIC*	*AIC*	*BIC*
Null model	604.34	639.56	672.31	707.53	689.01	727.75	665.82	701.04	686.41	721.62
Full model	586.63	628.89	664.08	706.34	677.32	723.10	647.69	689.95	674.30	716.56

Note: The null model excludes metacognitive strategies as a moderator; the full model includes it. AICs and BICs of the full models are consistently lower than the null models across measures of cognitive job resources, indicating a superior model fit of the full models.

References

Bliese, P. D., & Wang, M. (2020). Results provide information about cumulative probabilities of finding significance: Let's report this information. *Journal of Management, 46*(7), 1275-1288.

MacKinnon, D. (2008). *Introduction to statistical mediation analysis*. Erlbaum.

Nahhas, R. W. (2024). *Introduction to Regression Methods for Public Health Using R*. CRC Press.

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science, 381*(6654), 187-192.