Supplementary Materials to The Cognitive Cost of Going the Extra Mile: How Striving for Improvement Relates to Cognitive Performance

Table of Contents

Supplementary Material A: Experimental Materials in Studies 3 and 4

Experimental condition – Task proactivity

Oftentimes there are more efficient ways to complete our core tasks at work, and it would be possible to complete them in a way that saves time and effort. We would like to understand how you might improve the way in which your core tasks are done.

Please reflect on one core task in your job and suggest ways in which it could be done more efficiently. Please describe in detail how you would implement the change, outlining each action you would take to implement it. Please make sure to describe potential obstacles or negative consequences to implementing this change, and how you would address them.

Control condition – Routine task

Oftentimes there are standard ways to complete our core tasks at work and these ensure that tasks are completed properly. We would like to understand how you go about completing your core tasks following these routine procedures.

Please reflect on one core task that you have to routinely complete in your job. Please describe in detail how you perform this task, outlining each action you take to complete it. Please make sure to describe it so that we get a really clear sense of what it looks like when you perform this core task well.

Supplementary Material B: Means, Standard Deviations, Cronbach's Alpha, and Correlations (Study 3)

	Variable	*M*	*SD*	1	2	3	4	5	6	7	8
1	Task proactivity	2.68	1.23	(.94) (.89)	.15	.12	.03	−.01	−.01		
2	Mental effort (NASA-TLX subscale)	4.49	2.33	.46	(.91) (.89)	.69	−.09	−.00	−.07		
3	Mental effort (RSME)	60.53	31.26	.35	.67		−.05	.16	.10		
4	Routineness	3.65	0.90	−.20	−.33	−.27	(.77) (.58)	.11	.12		
5	Ego Depletion	1.88	0.81	.11	.11	.14	−.01	(.89) (.89)	.69		
6	Fatigue	1.94	0.95	.08	−.01	.09	.07	.74	(.93) (.92)		
	Correlations between routine task condition variables and between-subject variables										
7	Age	39.96	11.58	.01	.07	−.08	−.12	−.23	−.21		
8	Gender: Female	0.56	0.50	−.00	−.15	−.09	.05	.03	.02	−.04	
9	Gender: Male	0.43	0.50	.00	.15	.09	−.05	−.03	−.02	.04	−1.00
	Correlations between task proactivity condition variables and between-subject variables										
7	Age	39.96	11.58	−.01	.12	.01	−.10	−.26	−.22		
8	Gender: Female	0.56	0.50	−.11	−.19	−.15	.12	.04	.03	−.04	
9	Gender: Male	0.43	0.50	.11	.19	.15	−.12	−.04	−.03	.04	−1.00

Note. N = 318 for all variables apart from gender. N = 313 for correlations involving gender.
Correlations below the diagonal are related to the routine task condition (except for between-subject variables).
Correlations above the diagonal are related to the task proactivity condition (except for between-subject variables).
Cronbach's alphas are in parentheses. We present first alpha for the routine task condition followed by that of the task proactivity condition.
Correlations ≥ |.11| are $p < .05$ and correlations ≥ |.16| are significant at $p < .01$.

Supplementary Material C: Means, Standard Deviations, Cronbach's Alpha, and Correlations of Variables (Study 4)

	Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13
1	Age	41.41	12.04													
2	Gender: Male	0.47	0.50	−.02												
3	Gender: Female	0.52	0.50	.03	−.97											
4	Experimental condition	0.50	0.50	−.02	.01	−.04										
5	Task proactivity	3.38	1.27	.00	−.01	.00	.54	(.92)								
6	Mental effort (NASA-TLX subscale)	5.16	2.06	.17	.00	−.02	.11	.26	(.85)							
7	Mental effort (RSME)	55.05	26.63	.03	.01	−.04	.14	.19	.65							
8	Routineness	3.50	0.89	−.12	−.06	.07	−.36	−.32	−.19	−.16	(.69)					
9	Fatigue	2.16	1.07	−.36	−.10	.11	.00	−.05	−.09	.03	.07	(.94)				
10	Ego depletion	2.04	0.97	−.42	−.11	.10	−.02	−.03	−.05	.10	.05	.76	(.93)			
11	Self-control demands	2.16	0.76	−.04	.10	−.09	.07	.09	.21	.29	−.01	.12	.19	(.89)		
12	Self-control motivation	3.21	0.98	.01	.00	.03	.01	.08	.11	.07	.09	.04	.00	.31	(.93)	
13	Self-control effort	2.29	0.97	−.13	−.01	.02	−.01	.04	−.01	.11	.16	.27	.34	.42	.40	(.94)

Note. *N* = 319 for all variables apart from gender. *N* = 308 for correlations involving gender.
Cronbach's alphas are in parentheses.
Correlations ≥ | .11 | are $p < .05$ and correlations ≥ | .16 | are $p < .01$.

Supplementary Material D: Reverse Causality Experiment

While Study 3 and Study 4 strengthened our confidence in the causal relationship we propose between task proactivity and mental effort, we conducted an additional study exploring the potential reverse causal relationship between cognitive demands and task proactivity[1]. We investigated whether different levels of cognitive demands have an impact on the level of task proactivity. The procedures for this study were approved by the ESSEC Research Ethics Committee[2].

Participants, Design, and Procedure

Our between-subject experimental design included two conditions in which we manipulated cognitive demands by presenting participants either with the cognitively more demanding 2-back task or with the cognitively less demanding 1-back task. According to prior research (Sliwinski et al., 2006; Verhaeghen & Basak, 2005), in the 1-back task, the item remains in a state of immediate accessibility and recall is cognitively relatively effortless. In contrast, in the 2-back task, recall is cognitively effortful because of the high cognitive demand involved in switch attentional focus from one item to another. Each participant was randomly allocated to one of the two conditions. After the experimental manipulation, we assessed task proactivity as an outcome using the well-established situational judgement test by Bledow and Frese (2009). Participants reported also their perceived level of task proactivity.

We recruited participants based in the UK who were currently employed, via Prolific Academic. We required the survey to be completed on a computer or laptop to ensure that participants could install the software required for the n-back task. 823 participants successfully installed the software, performed the n-back task, and completed the survey. Of

[1] We thank an anonymous reviewer for this suggestion.
[2] Data from the project "Reverse causality: Cognitive demand and task proactivity" / IRB protocol number: 2023-18.

these, 13 participants erroneously completed part of the study multiple times, 66 participants spent more than 7 minutes on the n-back task, and 11 participants scored a negative d' (an established measure of sensitivity; Haatveit et al., 2010; Swets et al., 1961) (indicating they had inversed the instructions by hitting a key for non-targets instead of targets) in the n-back task. Finally, due to a technical issue, Prolific did not provide us with the demographics of 2 participants. We thus excluded these 92 participants from the analysis.

Our final sample consists of 731 participants, out of which 383 were randomly allocated to the non-demanding 1-back task while 348 were randomly allocated to the highly demanding 2-back task. Participants were on average 40.3 years old. Fifty-one percent of participants identified as male, 46.8% as female, and the remaining 17 participants identified as neither or preferred not to disclose their gender. Participants worked in a range of industries, such as business-related services (19.0%), education and arts (15.6%), and IT and scientific services (9.4%).

Measures

Unless otherwise indicated, we applied a Likert-type answering format, where response anchors ranged from 1 (*not at all*) to 5 (*to a great extent*).

Cognitive Performance. We measured cognitive performance following the same procedure as detailed in Study 1.

Situational Judgement Tests of Proactivity. We assessed proactivity using the well-established situational judgement test (SJT) by Bledow and Frese (2009). We chose three situations that represented task proactivity (items 2, 8, and 12) and obtained approval for their use from the authors. The situational judgement test presents participants with a series of work-related situations and requires them to indicate which of four response options to each situation they would most and least likely engage in. Based on the procedure by Bledow and Frese (2009), the most likely answers were rated -1 for low proactivity, 0 for medium proactivity, and +1 for high proactivity. The least likely answers were reverse scored. We computed the mean of the three items to obtain an overall SJT score.

Task Proactivity. We further assessed task proactivity with Griffin, Neal, and Parker's (2007) scale of 3 items, to establish that the SJT measure converged with the measure of task proactivity used in our other studies. We focused the question stem on the situations presented in the SJT: "In the response options for the situations you selected, to what extent do you feel your response meant that you...". An example item is "initiated better ways of doing my core tasks" (α = .88).

Results

Table D.1 shows the descriptive statistics across experimental conditions. Overall sample statistics and correlations are shown in Table D.2.

Table D.1: Descriptive Statistics of Variables by Experimental Condition

	Cognitive Performance		*SJT of Proactivity*		*Task Proactivity*	
	M	*SD*	*M*	*SD*	*M*	*SD*
Experimental condition						
Low cognitive demand (1-back)	5.12	1.17	−0.52	0.85	2.98	0.96
High cognitive demand (2-back)	2.17	2.02	−0.55	0.85	2.96	1.03

Note: $N = 731$.

Table D.2: Means, Standard Deviations, Cronbach's alpha, and Correlations

	Variable	*M*	*SD*	1	2	3	4	5	6	7
1	Age	40.27	17.69							
2	Gender: Female	0.48	0.50	−0.02						
3	Gender: Male	0.52	0.50	0.02	−1.00					
4	Experimental condition	1.48	0.50	0.08	0.03	−0.03				
5	Cognitive performance	3.72	2.20	−.0.36	−0.04	0.04	−0.67			
6	SJT of proactivity	−0.54	0.85	−0.03	−0.10	0.10	−0.02	−0.08		
7	Task proactivity	2.97	1.00	0.00	−0.09	0.09	−0.01	−0.13	0.43	(0.88)

Note. $N = 731$ for all variables apart from gender. $N = 714$ for correlations involving gender. Cronbach's alpha is in parentheses.
Correlations ≥ |.11| are significant at $p < .05$ and correlations ≥ |.16| are significant at $p < .01$.

We performed t-tests to compare cognitive performance on the n-back tasks, SJT of proactivity, and task proactivity scores in our two experimental conditions (cognitively less demanding condition with 1-back versus cognitively more demanding condition with 2-back). Participants in the 1-back condition achieved higher levels of cognitive performance (M = 5.12, SD = 1.17) compared to those in the 2-back condition (M = 2.17, SD = 2.02; $t(729) = 24.361$, $p < .001$), providing support for the more demanding nature of the 2-back task.

There was no significant difference in SJT scores between the 1-back condition (M = −0.52, SD = 0.85) and the 2-back condition (M = −0.55, SD = 0.85; $t(729) = 0.581$, $p = .281$ (one-tailed)). Similarly, there was no significant difference in task proactivity between the 1-back condition (M = 2.98, SD = 0.96) and the 2-back condition (M = 2.96, SD = 1.03; $t(729) = 0.271$, $p = .393$). We thus found no support for the causal relationship between cognitive demands and task proactivity.

References

Bledow, R., & Frese, M. (2009). A situational judgment test of personal initiative and its relationship to performance. *Personnel Psychology*, *62(2)*, 229-258. https://doi.org/10.1111/j.1744-6570.2009.01137.x

Griffin, M. A., Neal, A., & Parker, S. K. (2007). A new model of work role performance: Positive behavior in uncertain and interdependent contexts. *Academy of Management Journal*, *50(2)*, 327-347. https://doi.org/10.5465/amj.2007.24634438

Haatveit, B. C., Sundet, K., Hugdahl, K., Ueland, T., Melle, I., & Andreassen, O. A. (2010). The validity of d prime as a working memory index: Results from the “Bergen n-back task”. *Journal of Clinical and Experimental Neuropsychology*, *32(8)*, 871-880. https://doi.org/10.1080/13803391003596421

Sliwinski, M. J., Smyth, J. M., Hofer, S. M., & Stawski, R. S. (2006). Intraindividual coupling of daily stress and cognition. *Psychology and Aging*, *21(3)*, 545-557. https://doi.org/10.1037/0882-7974.21.3.545

Swets, J. A., Tanner Jr, W. P., & Birdsall, T. G. (1961). Decision processes in perception. *Psychological Review*, *68(5)*, 301-340. https://doi.org/10.1037/h0040547

Verhaeghen, P., & Basak, C. (2005). Ageing and switching of the focus of attention in working memory: Results from a modified N-Back task. *The Quarterly Journal of Experimental Psychology Section A*, *58(1)*, 134-154. https://doi.org/10.1080/02724980443000241

Supplementary Material E: Measure Instructions and Adapted Items for Studies 3 and 4, the Laboratory Experiment (see Supplementary Material F), and the Online Experiment (see Supplementary Material G)

Mental Effort – adapted from a Subscale of the NASA-TLX (Hart & Staveland, 1988)

Think about the moment when you wrote down your core task *(routine task condition)*

Think about the moment when you wrote down the changes you could implement *(task proactivity condition)*

1. How much mental activity was required of you in terms of thinking hard, reflecting on pros and cons?
2. How much mental activity was required of you in terms of anticipating consequences, making decisions?
3. How much mental activity was required of you in terms of searching for ideas, looking for solutions?

Mental Effort – Rating Scale Mental Effort (RSME) adapted from Zijlstra (1995)

How hard did you have to work mentally in order to provide the information required for the previous description?

Routineness adapted from Chung & Jackson (2013)

On the previous page you described one of the core tasks you need to accomplish well at work. To what extent does the following apply to what you described? Completing my core task at work using the standard procedures means that… *(routine task condition)*

On the previous page you described changes you might implement to improve your way of doing one of your core tasks at work. To what extent does the following apply to what you described? Initiating better ways of doing my core task at work means that… *(task proactivity condition)*

1. The way of accomplishing my task will be very routine
2. The methods I will follow in my work are about the same as the ones prescribed at work
3. I will feel like I am doing the same thing over and over again

Supplementary Material F: Laboratory Experiment

While Study 3 and Study 4 strengthened our confidence in the causal relationship we propose between task proactivity and mental effort, we conducted a lab experiment[3] investigating whether increased mental effort and reduced routineness mediate the effects of task proactivity on cognitive performance. This experiment was pre-registered on Aspredicted (https://aspredicted.org/gk39w.pdf).

Participants, Design, and Procedure

Our between-subject design included two conditions, using the same manipulation of task proactivity as in Studies 3 and 4: a) a routine task condition in which participants were asked to reflect on a core task that they have to routinely complete in their job and describe it in detail, and b) a task proactivity condition in which participants were asked to reflect on a core task in their job, suggest ways in which it could be done more efficiently, and describe in detail how they would implement this change. Participants were randomly allocated to one of the two conditions. Before the experimental condition (routine task condition or task proactivity condition), participants reported their ego depletion and subjective fatigue levels and completed a 2-back task to measure their pre-test cognitive performance. After the experimental condition, participants reported their perceived level of task proactivity, experienced mental effort, routineness, ego depletion, and subjective fatigue, and then completed the 2-back task for a second time to measure their post-test cognitive performance.

We recruited 62 participants at the third author's institution. All of them were students with current or past professional experience. As pre-registered, we excluded 2 participants who spent less than 3 minutes on the writing tasks[4]. We also excluded 4 participants who

[3] This experiment was conducted in the psychology department of Potsdam University. Following the criteria of the German Research Foundation, approval of the ethics committee was not required.

[4] When these exclusion criteria are not applied, our findings remain largely large unchanged. The effect of the experimental conditions on task proactivity remains non-significant ($t(60) = -0.737$, $p = .232$). There were no significant differences for the RSME measure of mental effort ($t(60) = -1.349$, $p = .091$) and routineness ($t(60) = 1.563$, $p = .062$). Repeated-measures ANOVAs with pre- and post-test measures, showed no significant

spent more than 5 minutes on the second n-back task. Finally, we excluded 2 participants who got a d’ (measure of sensitivity) of zero (chance performance) or less (inversing the instructions by hitting a key for non-targets instead of targets) in the n-back task. No participant failed to respond correctly to any of our instructed response items. We thus excluded a total of 8 participants from the analysis.

Our final sample consists of 54 participants, out of which 29 were randomly allocated to the task proactivity condition while 25 were randomly allocated to the routine task condition. Participants were on average 23.5 years old. Thirteen percent of participants identified as male, 85.2% as female, and the remaining participant identified as neither. Participants reported working in a range of industries, such as food and hospitality (22.2%), education and arts (22.2%), and commercial trade (14.8%).

Measures

Unless mentioned otherwise, we applied the same measures as in Studies 3 and 4, focusing question stems on the experimental manipulation through the writing tasks (e.g., “To what extent does the following apply to what you described?”).

Cognitive Performance. We collected cognitive performance using the 2-back task before and after the experimental manipulation. We used the same task and measure as in Studies 1 and 2.

Task Proactivity ($\alpha = .90$), mental effort (NASA-TLX subscale measure: $\alpha = .85$), routineness ($\alpha = .75$.), ego depletion ($\alpha_{pre\text{-}test} = .88$; $\alpha_{post\text{-}test} = .90$) and subjective fatigue ($\alpha_{pre\text{-}test} = .86$; $\alpha_{post\text{-}test} = .80$) were assessed with the same measures as in Study 3.

Results

differences between the task proactivity condition and the routine task condition in terms of cognitive performance ($F(1, 60) = 0.322$, $p = .573$, partial $\eta^2 = .005$), ego depletion ($F(1, 60) = 0.235$, $p = .630$, partial $\eta^2 = .004$), and subjective fatigue ($F(1, 60) = 0.238$, $p = .628$, partial $\eta^2 = .004$). Mental effort measured with the NASA-TLX subscale was significantly higher in the task proactivity group ($M = 4.70$, $SD = 1.94$) than in the control group ($M = 3.78$, $SD = 2.17$); $t(60) = -1.761$, $p = .042$).

We performed t-tests and repeated-measures ANOVAs to evaluate the effect of our experimental conditions (routine task condition versus task proactivity condition) on task proactivity (manipulation check), experienced mental effort and routineness (proposed mechanisms), ego depletion and subjective fatigue (competing mechanisms), and cognitive performance (dependent variable). Table F.1 shows the descriptive statistics across experimental conditions. Overall sample statistics and correlations are shown in Table F.2.

Table F.1: Descriptive Statistics of Variables by Experimental Condition

	Task proactivity condition		Routine task condition	
	M	*SD*	*M*	*SD*
Task proactivity	3.57	0.93	3.36	1.26
Mental effort (NASA-TLX subscale)	4.56	1.97	3.80	2.15
Mental effort (RSME)	54.83	24.59	49.20	28.71
Routineness	3.44	0.67	3.67	0.94
Ego depletion – pre-test	2.36	0.95	2.22	0.79
Ego depletion – post-test	2.21	0.93	2.22	0.76
Subjective fatigue – pre-test	2.22	0.96	2.19	0.85
Subjective fatigue – post-test	2.01	0.79	2.07	0.72
Cognitive performance – pre-test	2.83	1.60	2.69	1.77
Cognitive performance – post-test	3.26	2.04	3.33	1.93

Note: *N* = 54.

Table F.2: Means, Standard Deviations, Cronbach's alpha, and Correlations

	Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13
1	Age	23.48	5.34													
2	Gender: Female	0.87	0.34	–0.12												
3	Gender: Male	0.13	0.34	0.12	–1.00											
4	Experimental condition	0.54	0.50	0.34	–0.03	0.03										
5	Task proactivity	3.48	1.09	–0.11	–0.26	0.26	0.10	(.90)								
6	Mental effort (NASA-TLX subscale)	4.21	2.07	0.34	–0.34	0.34	0.19	0.23	(.85)							
7	Mental effort (RSME)	52.22	26.47	0.35	–0.35	0.35	0.11	0.10	0.66							
8	Routineness	3.54	0.81	–0.39	0.10	–0.10	–0.14	0.08	–0.39	–0.25	(.75)					
9	Ego depletion – pre-test	2.30	0.87	–0.02	–0.04	0.04	0.09	–0.19	0.18	0.05	–0.30	(.88)				
10	Ego depletion – post-test	2.21	0.84	0.03	–0.18	0.18	–0.01	–0.09	0.36	0.31	–0.27	0.80	(.90)			
11	Subjective fatigue – pre-test	2.21	0.91	–0.04	–0.00	0.00	0.02	–0.08	0.27	0.22	–0.27	0.69	0.68	(.86)		
12	Subjective fatigue – post-test	2.04	0.75	–0.15	–0.14	0.14	–0.04	0.04	0.27	0.22	–0.24	0.63	0.76	0.88	(.80)	
13	Cognitive performance – pre-test	2.77	1.67	–0.12	–0.17	0.17	0.04	0.28	0.22	0.18	–0.08	0.02	0.13	0.11	0.16	
14	Cognitive performance – post-test	3.30	1.97	–0.16	–0.04	0.04	–0.02	0.10	–0.09	0.08	–0.02	–0.02	–0.00	0.09	0.07	0.71

Note: $N = 54$ for all variables apart from gender. $N = 53$ for correlations involving gender.
Cronbach's alphas are in parentheses.
Correlations $\geq |.11|$ are significant at $p < .05$ and correlations $\geq |.16|$ are significant at $p < .01$.

The effect of the experimental conditions on task proactivity was not significant ($t(52) = -0.718$, $p = .238$; see Table F.1), suggesting that in this student sample our manipulation of task proactivity was not successful. There were also no significant differences between experimental conditions in mental effort (NASA-TLX subscale measure: $t(52) = -1.361$, $p = .090$; RSME measure: $t(52) = -0.776$, $p = .221$) and routineness ($t(52) = 1.045$, $p = .150$)

Repeated-measures ANOVAs with pre- and post-test measures, showed no significant differences between the task proactivity condition and the routine task condition in terms of cognitive performance ($F(1, 52) = 0.270$, $p = .605$, partial $\eta^2 = .005$), ego depletion ($F(1, 52) = 1.153$, $p = .288$, partial $\eta^2 = .022$), and subjective fatigue ($F(1, 52) = 0.630$, $p = .431$, partial $\eta^2 = .012$).

Discussion

Together, these results suggest that in this laboratory experiment with a student sample our experimental manipulation has not been successful. While our manipulation of task proactivity had been effective in Studies 3 and 4 with samples of more experienced professionals, it may be that the limited work experience of our student participants restricted their ability to envision meaningful instances of task proactivity.

However, an additional post-hoc regression analysis (Table F.3) shows that, controlling for pre-test cognitive performance and experimental condition, NASA-TLX mental effort was negatively associated to post-test cognitive performance ($\gamma = -0.24$, $SE = 0.10$, $p = .018$; see Table F.3). While we did thus not find any differences in cognitive performance between experimental conditions, we found some support for our argument that higher levels of mental effort would result in lower cognitive performance.

Table F.3: Means, Standard Deviations, Cronbach's alpha, and Correlations

	DV: post-test cognitive performance	
	γ	*SE*
Intercept	1.89 **	(0.62)
Pre-test cognitive performance	0.90 ***	(0.11)
Condition (1 = task proactivity; 0 = routine task)	−0.00	(0.38)
Mental effort	−0.24 *	(0.10)
Ego depletion – post-test	−0.03	(0.24)

N = 54; adjusted R^2 = .538.
* *p < .05;* ** *p < .01;* *** *p < .001 (two-tailed)*

Supplementary Material G: Online Experiment Including Cognitive Performance

Study 3 and Study 4 established the causal relationship between task proactivity and mental effort, yet they did not test the effects of task proactivity on cognitive performance. We thus conducted a further online experiment including the n-back task[5]. The procedures for this study were approved by the ESSEC Research Ethics Committee[6]. This experiment was pre-registered on Aspredicted (https://aspredicted.org/kk9er.pdf).

Participants, Design, and Procedure

We used the same between-subject design and manipulation of task proactivity as in the laboratory experiment described in Supplementary Material F. As in these prior studies, participants were randomly allocated to a) the routine task condition or b) the task proactivity condition. Participants were randomly allocated to one of the two conditions. Replicating the design of the laboratory experiment, participants reported their ego depletion and subjective fatigue levels and completed a 2-back task to measure their pre-test cognitive performance prior to allocation to the experimental manipulation, and following the experimental manipulation reported their perceived level of task proactivity, experienced mental effort, routineness, ego depletion, and subjective fatigue, and then completed the 2-back task for a second time to measure their post-test cognitive performance.

We recruited 692 participants based in the UK. For 19 participants no data was recorded in the n-back task, resulting in an initial sample of 673. As pre-registered, we excluded 138 participants who spent less than 3 minutes on the writing tasks, 19 participants who spent more than 5 minutes on the second n-back task, and 6 participants who got a d' (measure of sensitivity) of zero (chance performance) or less (inversing the instructions by hitting a key for non-targets instead of targets) in the n-back task. No participant failed to

[5] We are grateful to an anonymous reviewer for this suggestion.
[6] Data from project "Experiment for R2: Task proactivity and working memory performance" / IRB protocol number: 2023-32.

respond correctly to any of our instructed response items. We also excluded 4 participants who encountered technical problems when starting the n-back task and therefore completed part of the study multiple times[7]. We thus excluded a total of 167 participants from the analysis. This resulted in a final sample of 506 participants, out of which 257 were randomly allocated to the task proactivity condition and 249 were randomly allocated to the routine task condition. Participants were on average 39.1 years old. Two hundred sixty-seven participants identified as male, 238 as female, and the remaining participant identified as neither. Participants reported working in a range of industries, the most frequent being business-related services (17.6%), IT and scientific services (13.2%), and healthcare (12.5%).

Measures

We applied the same measures as in the laboratory experiment described in Supplementary Material F. Reliabilities can be found in Table G2.

Results

As preregistered, we performed t-tests to evaluate the effect of our experimental conditions (routine task condition versus task proactivity condition) on task proactivity (manipulation check), experienced mental effort and routineness (proposed mechanisms), and cognitive performance (dependent variable). For exploratory purposes we also tested the effects on the competing theoretical mechanisms of fatigue and ego depletion. Descriptive statistics across experimental conditions are show in Table G.1. Table G.2 shows the overall sample statistics and correlations.

[7] Twenty-seven participants dropped out after going through the pre-test cognitive performance but before getting to the manipulation. After the manipulation, 8 participants dropped out from the control condition (7 did not get to the post-test cognitive performance and 1 did the post-test but did not finish the study) and 6 participants dropped out from the experimental condition (6 did not get to the post-test cognitive performance).

Table G.1: Descriptive Statistics of Variables by Experimental Condition

	Task proactivity condition		Routine task condition	
	M	*SD*	*M*	*SD*
Task proactivity	3.78	0.97	2.94	1.31
Mental effort (NASA-TLX subscale)	5.29	2.06	4.83	2.39
Mental effort (RSME)	72.76	27.78	61.77	30.57
Routineness	3.04	0.82	3.58	0.90
Ego depletion – pre-test	2.11	0.99	2.06	0.95
Ego depletion – post-test	2.02	1.01	1.90	0.92
Subjective fatigue – pre-test	2.05	0.92	1.96	0.91
Subjective fatigue – post-test	1.97	0.94	1.89	0.90
Cognitive performance – pre-test	2.29	1.93	2.52	1.98
Cognitive performance – post-test	2.95	2.12	2.95	2.14

Note: $N = 506$.

Table G.2: Means, Standard Deviations, Cronbach's alpha, and Correlations

	Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	Age	39.14	11.01															
2	Gender: Female	0.47	0.50	-.02														
3	Gender: Male	0.53	0.50	.02	-1.00													
4	Experimental condition	0.51	0.50	.03	.00	.00												
5	Task proactivity	3.36	1.22	.01	-.03	.03	.34	(.94)										
6	Mental effort (NASA-TLX subscale)	5.06	2.24	.08	-.06	.06	.10	.29	(.88)									
7	Mental effort (RSME)	67.35	29.67	.04	.00	.00	.19	.20	.69									
8	Routineness	3.31	0.90	-.11	.15	-.16	-.30	-.16	-.22	-.17	(.65)							
9	Ego depletion – pre-test	2.08	0.97	-.16	.07	-.08	.03	.02	.00	.08	.05	(.92)						
10	Ego depletion – post-test	1.97	0.97	-.19	.06	-.07	.06	.00	.10	.20	.04	.81	(.93)					
11	Subjective fatigue – pre-test	2.00	0.92	-.21	.09	-.10	.05	.02	.01	.08	.09	.74	.73	(.92)				
12	Subjective fatigue – post-test	1.93	0.92	-.22	.05	-.06	.04	-.01	.05	.14	.08	.67	.82	.84	(.92)			
13	Cognitive performance – pre-test	2.40	1.95	-.03	-.03	.03	-.06	-.03	-.17	-.20	.03	-.02	-.09	.02	-.05			
14	Cognitive performance – post-test	2.95	2.13	.08	-.04	.04	.00	-.04	-.14	-.19	-.01	-.06	-.14	-.01	-.10	.73		
15	Job complexity	2.13	1.02	.00	.09	-.08	.03	-.11	-.11	-.07	.24	-.01	-.02	.01	.00	.00	.02	(.88)

Note: $N = 506$.
Cronbach's alphas are in parentheses.
Correlations $\geq |.10|$ are significant at $p < .05$ and correlations $\geq |.14|$ are significant at $p < .01$.
Experimental condition is coded as 0 = routine task condition, 1 = task proactivity condition

The effect of the experimental conditions on task proactivity was significant ($t(527) = 8.183$, $p < .001$ (one-tailed); see Table G.2), suggesting that our manipulation of task proactivity was successful. As in Studies 3 and 4, mental effort was significantly higher in the task proactivity condition than in the routine task condition (NASA-TLX subscale measure: $t(504) = 2.284$, $p = .005$ (one-tailed); RSME measure: $t(504) = 4.237$, $p < .001$ (one-tailed)), and routineness was significantly lower (($t(504) = -6.977$, $p < .001$ (one-tailed)). As in previous studies, the two experimental conditions did not differ in post-test fatigue ($t(504) = 0.968$, $p = .17$ (one-tailed)) or ego depletion (($t(504) = 1.409$, $p = .08$ (one-tailed). There were no differences in conditions in post-test cognitive performance ($t(504) = -0.013$, $p = .495$ (one-tailed)). Our hypothesis that simulating task proactivity would result in decreased cognitive performance compared to simulating engaging in a routine task was thus not supported.

Discussion

While we were able to once again replicate the effects of task proactivity on mental effort and routineness, we found no support for our hypothesis regarding the effect of simulating task proactivity on cognitive performance in this pre-registered study.

Exploratory post-hoc analyses

For exploratory purposes we had included job complexity as a potential control variable, as mentioned in the pre-registration. Job complexity refers to the extent to which a person's job is mentally demanding. The job complexity of participants in our sample was low ($M = 2.13$, $SD = 1.02$), compared to $M = 3.85$ ($SD = 0.73$), which is the level of job complexity reported in the publication presenting the development of the measure we used (Morgeson & Humphrey, 2006). To obtain a more robust benchmark to compare the job complexity of the present sample to, we searched for further publications that used Morgeson and Humphrey's (2006) measure of job complexity. We conducted a cited reference search (focusing primarily on publications with "complexity" in title) and consulted the website dedicated to the measure (Morgeson, 2021). This preliminary search resulted in a total of 12 samples (*n* = 5,062). Sample weighed mean was *M* = 3.75 (*SD* = 0.98). Comparing this to the mean of the present sample (*M* = 2.15, *SD* = 1.03) made us suspect that many of our respondents worked in low-complexity jobs that do not involve the use of high-level skills and that are mentally not very demanding (Humphrey et al., 2007). It may be that in such low-complexity jobs which involve simple and uncomplicated tasks the mental simulation of task proactivity our experimental manipulation required was cognitively not so demanding that it would have impacted subsequent cognitive performance. To explore this possibility, we split the sample in two at the scale midpoint of the job complexity scale[8] and conducted a 2x2 ANOVA. The resulting "typical job complexity group" contained 116 participants, with an average job complexity score of *M* = 3.66 (*SD* = 0.59), and the "low job complexity group" contained the remaining 390 participants with an average job complexity score of 1.68 (*SD* = 0.59). The main effects of job complexity or experimental condition were not significant ($F(1, 502) = 0.019$, $p = .164$, and $F(1, 502) = 1.964$, $p = .164$, respectively) but their interaction was significant $F(1, 502) = 6.541$, $p = .011$). Tests of simple main effects showed that cognitive performance was significantly lower in the task proactivity condition than in the routine task condition in the typical job complexity group ($p = .025$). There was no significant difference between conditions in the low job complexity group ($p = .225$). Thus, there is preliminary support for our hypothesis when focusing on a subsample that has a typical level of job complexity. These exploratory post-hoc analyses thus provide some

[8] We chose this approach over treating job complexity as a continuous co-variate because it was strongly skewed: 21.7% of participants reported the lowest possible level of complexity, and over 55% of participants had a score of 2 or lower. We opted to split the sample at the scale midpoint rather than the median because a median split would have resulted in a typical job complexity group in which the average job complexity was at 2.92 (*SD* = 0.79) which would have been still lower than in comparable prior studies (*M* = 3.75, *SD* = 0.98, *n* = 5,062). Using the midpoint instead resulted in a mean (*M* = 3.66, *SD* = 0.59) that was more similar to what has been typically observed.

initial support for the notion that in a sample with a typical level of job complexity the mental simulation of task proactivity may be cognitively demanding. Further research in samples with job complexity levels more typical of the general working population is needed to explore the relationship between simulated task proactivity and cognitive performance.

We also explored whether controlling for pre-test cognitive performance would allow us to detect a significant effect. We thus included pre-test cognitive performance as a covariate in an ANCOVA predicting post-test cognitive performance. Controlling for pre-test performance, there were no differences in conditions in post-test cognitive performance ($F(1,526) = 2.520$, $p = .113$). Our hypothesis was not supported.

In summary, in this study, our pre-registered test of our hypothesis that simulating task proactivity would result in lower cognitive performance was not significant. Exploratory post-hoc analyses hint at the possibility that a sample with job complexity levels typical of the general working population may be required to be able to detect an effect.

A further possible explanation for why we found an association between task proactivity during the working day and end-of-day cognitive performance in our field studies, including when controlling for beginning-of-day cognitive performance, but no effect of task proactivity on subsequent cognitive performance in this experiment may be that our experimental manipulation was only able to simulate the envisioning stage of the proactive behavior process (Bindl et al., 2012). Our proactivity manipulation required participants to imagine how they could engage in task proactivity, but it did not require them to engage in the enacting and reflecting stages. When proactively making changes to their core task individuals do not only identify and imagine possible improvements, as they did in our experiment, but also take action to implement them, deal with unpredicted consequences, observe outcomes, and make further adjustments (Bindl et al., 2012). While we consistently found that even merely envisioning task proactivity is associated with increased reported mental effort, it may be that in Studies 1 and 2 the taxing demands of the additional stages of task proactivity accumulated over the course of the day. Further research is needed to explore this possibility, for example, in an experimental experience sampling study.

References

Bindl, U. K., Parker, S. K., Totterdell, P., & Hagger-Johnson, G. (2012). Fuel of the self-starter: How mood relates to proactive goal regulation. *Journal of Applied Psychology*, *97*(1), 134-150. https://doi.org/10.1037/a0024368

Humphrey, S. E., Nahrgang, J. D., & Morgeson, F. P. (2007). Integrating motivational, social, and contextual work design features: A meta-analytic summary and theoretical extension of the work design literature. *Journal of Applied Psychology*, *92*(5), 1332-1356. https://doi.org/10.1037/0021-9010.92.5.1332

Morgeson, F. P. (2021). *WDQ: The work design questionnaire*. Retrieved 11/24/23 from http://www.morgeson.com/wdq.html

Morgeson, F. P., & Humphrey, S. E. (2006). The Work Design Questionnaire (WDQ): Developing and validating a comprehensive measure for assessing job design and the nature of work. *Journal of Applied Psychology*, *91*(6), 1321-1339. https://doi.org/10.1037/0021-9010.91.6.1321