

Interaction Techniques that Encourage Longer Prompts Can Improve Psychological Ownership when Writing with AI

Nikhita Joshi*

Cheriton School of Computer Science
University of Waterloo
Waterloo, Ontario, Canada
nvjoshi@uwaterloo.ca

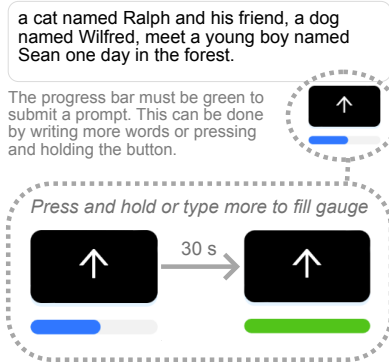
Daniel Vogel

Cheriton School of Computer Science
University of Waterloo
Waterloo, Ontario, Canada
dvogel@uwaterloo.ca

(a) *motor cost*: pump slider



(b) *time cost*: press and hold button



(c) receiving suggestions from AI

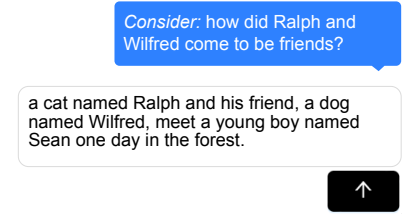


Figure 1: Augmented prompt entry techniques that (a) increase the motor cost of submitting a prompt by making the user ‘pump’ a gauge by moving a slider; (b) increase the temporal cost by requiring users to press and hold the button to fill a gauge; and (c) provide AI-generated suggestions for how the user’s prompt can be expanded.

Abstract

Writing longer prompts for an AI assistant to generate a story increases psychological ownership, a user’s feeling that the writing belongs to them. To encourage users to write longer prompts, we evaluated two interaction techniques that modify the prompt entry interface of chat-based generative AI assistants: pressing and holding the prompt submission button, and continuously moving a slider up and down when submitting a short prompt. A within-subjects experiment investigated the effects of such techniques on prompt length and psychological ownership, and results showed that these techniques increased prompt length and led to higher psychological ownership than baseline techniques. A second experiment further augmented these techniques by showing AI-generated suggestions for how the prompts could be expanded. This further increased prompt length, but did not lead to improvements in psychological ownership. Our results show that simple interface modifications like these can elicit more writing from users and improve psychological ownership.

CCS Concepts

• **Human-centered computing** → **Empirical studies in HCI**; **Interaction techniques**.

*Also with LISN, Université Paris-Saclay, CNRS, Inria.

Keywords

generative AI, large language models, interaction techniques, controlled experiments

1 Introduction

Psychological ownership is a concept in psychology that describes possessive feelings people have towards anything, for example, ideas, places, objects, and even creations, regardless of legal ownership [33, 34]. Writers often feel psychological ownership for written text [2, 5, 16, 32, 39], likely because writing requires an investment of time, energy, and sense of self, all of which are important for psychological ownership [33, 34, 40]. Losing too much psychological ownership can happen when writing collaboratively [2, 5, 16, 39], which can lead to negative outcomes like territorial behaviours [24, 39] and can even deter future collaborations [5].

More recently, writers have been ‘collaborating’ with generative AI tools like ChatGPT. Research has shown that psychological ownership is also important for human-AI collaborative writing systems [25, 35], leading to increased interest within the HCI community in how using these tools affects psychological ownership (e.g., [12, 27, 29]). Notably, Joshi and Vogel [19] explored the role of *prompt length* on psychological ownership when writing short stories, and found that writing longer prompts increased psychological ownership. Participants were encouraged to provide more details about the story plot; however, this was primarily achieved through a strict word minimum on the prompt length. In practice, users desire more flexibility. One way to encourage longer prompts without preventing shorter prompts is to create interaction techniques that

purposely introduce *friction* when users write short prompts by introducing *costs*. As an example, they show that requiring the user to press and hold the prompt submission button, a time cost, can encourage users to write longer prompts, but they did not evaluate the effects of this technique on psychological ownership.

There are many other ways to introduce friction into the prompt submission interface to encourage users to write longer prompts for increased psychological ownership. Inspired by their technique, we propose augmenting the prompt entry interface so users have to fill a ‘gauge’ by writing more words, or by repeatedly moving a slider up and down (Figure 1). Much like Joshi and Vogel’s time delay technique, this provides users with a choice: they can either write a longer prompt, or exert physical effort to submit a shorter one. We evaluated the effect of both augmented prompt entry interfaces on prompt length and psychological ownership. Twenty nine participants wrote prompts to generate short stories for a generative AI writing assistant and answered subjective questions about psychological ownership and task workload. When compared to a baseline interface with no augmentations, participants felt more mental demand, physical demand, effort, and frustration, but wrote significantly longer prompts and felt more psychological ownership toward the resulting stories. To encourage participants to generate more ideas [36] and write even longer prompts, we further augmented these techniques to display AI-generated suggestions every 8 seconds, triggered after one second of inactivity. A follow up study with 30 participants showed that this led to significantly longer prompts for the augmented prompt entry conditions, but did not further improve psychological ownership. Our work contributes:

- a new interaction technique for increasing prompt length that increases the motor cost of submitting a short prompt;
- empirical evidence that shows this technique improves psychological ownership and confirms the effectiveness of Joshi and Vogel’s technique;
- empirical evidence that shows that receiving AI-generated suggestions increases prompt length, but does not improve psychological ownership.

2 Background and Related Work

Our work relates to previous work examining psychological ownership when writing with AI, and previous work that introduces friction into user interfaces.

2.1 Psychological Ownership when Writing with AI

Given the increase of writing tools with AI capabilities, researchers have been advocating for AI writing systems that are designed [25] and evaluated [35] in a way that prioritizes psychological ownership. Some prior work has explored the effects of the role of AI writing assistants on psychological ownership. Biermann et al. [1] presented writers with mock-up user interfaces that utilized AI assistance in different ways when writing stories (e.g., using AI to generate character dialog, or to elaborate on the scene’s setting), and found that many designs were not well-received as they reduced psychological ownership. Draxler et al. [12] explored using an AI assistant for writing postcards. They found that when the AI assistant wrote the entire message, participants did not feel much

psychological ownership, yet did not attribute authorship to AI (i.e., an AI Ghostwriter Effect). Li et al. [29] showed that when AI acted as the writer, psychological ownership was reduced for participants when compared to the AI acting as an editor. However, they were willing to accept lower remuneration in exchange for receiving writing assistance from the AI assistant, signifying its high value.

The type and quantity of AI suggestions during the writing process can affect psychological ownership. Lehmann et al. [27] showed that manually integrating AI-generated suggestions into text led to higher feelings of psychological ownership than those that were automatically integrated. Dhillon et al. [10] showed that when writers received suggestions, psychological ownership decreased, yet this effect was more pronounced for longer suggestions (paragraphs) than shorter ones (sentences). Lee et al. [26] speculate that receiving more suggestions from AI reduces how much the human writes, leading to lower psychological ownership.

Joshi and Vogel [19] were the first to provide empirical evidence that the ratio of human-to-AI-generated text can influence psychological ownership. Participants wrote prompts of pre-defined lengths (ranging from 3 to 150 words), which were enforced through word minimums and maximums to generate short stories. As participants wrote longer prompts, they felt more psychological ownership towards the resulting stories. This was likely because participants included more details about the story in their prompts, which required additional thought and effort.

2.2 Friction in User Interfaces

Though it may feel counterintuitive to purposely make the act of submitting a prompt more difficult, prior work shows that this can be beneficial. Notably, friction in design can encourage users to move from fast, automatic, and intuitive “System 1” thinking to slow, deliberate, and reflective “System 2” thinking that is more conducive to increased attention and analytical thinking [3, 4, 8, 20], which can encourage changes in behaviour [4, 8].

Friction can be introduced into user interfaces in many ways, for example, by increasing temporal costs. Wang et al. [41] showed that time delays can encourage increased reflection before posting to social media. Lewis and Vogel [28] used time delays to encourage users to transition to expert-mode menu selection techniques [28]. When using an AI assistant, Bućinca et al. [3] required users to wait before displaying AI-generated suggestions, which encouraged more analytical thinking about the output and reduced over-reliance [3]. Joshi and Vogel [19] proposed augmenting the prompt submission button so users had to press and hold it for a few seconds when submitting a shorter prompt. This effectively increased prompt length, but they did not verify the effect on psychological ownership.

Beyond increasing the temporal costs of submitting a shorter prompt, there are many other ways users could be encouraged to write longer prompts, and prior work has shown the effectiveness of other costs in non-AI contexts. For example, increasing the perception cost of reading text can reduce errors when transcribing it [38], and increasing the perception and motor costs of viewing an interface (i.e., by “brushing” away “frost” on the screen) can improve spatial tasks [7].

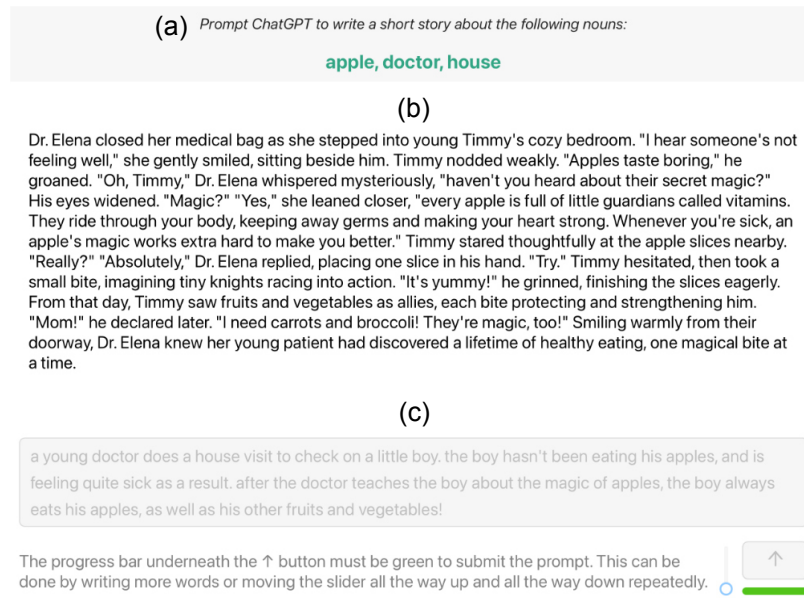


Figure 2: Experimental interface (shown with MOTOR): (a) instructions and nouns the story had to be about; (b) the AI-generated short story; and (c) the augmented prompt entry interface.

Overall, psychological ownership is important when writing with AI, and factors, including prompt length, can affect it. However, previous work has not evaluated the effect on psychological ownership for techniques designed to encourage writing longer prompts by introducing friction through costs.

3 Augmented Prompt Entry Techniques

We evaluate the effectiveness of two different augmented prompt entry techniques for increasing prompt length and psychological ownership: one new technique that increases the motor cost of submitting a short prompt, and another from Joshi and Vogel that increases the temporal cost of submitting a short prompt [19].¹ Both techniques are based on the idea of having to fill a 'gauge' to submit a prompt; an idea that proved to be successful among several other ideas originally tested by Joshi and Vogel. We compare both of these to a baseline condition, where the user can press the prompt submission button normally, whenever they wish, and without any additional costs.

3.1 Motor Cost

Our new interaction technique introduces friction by increasing the *motor cost* of submitting a short prompt. Below the prompt submission button is a small progress bar that represents a 'gauge' that the user must fill. This can be done by writing more words in the prompt submission text box, or by 'pumping' it full by moving a 45 pixel tall vertical slider up and down repeatedly (Figure 1a).

The progress bar width is the ratio of the number of words in the prompt to the *word minimum*. We use 150 words to align with Joshi and Vogel's previous experiment [19]. When the user moves the slider up and down, the progress bar width increases. The number of movements the user must perform to fill the gauge is calculated

by mapping the ratio of words typed to a *movement maximum* (225 movements, which was determined through a pilot experiment; see Appendix A for more detail). For example, if the user types 15 words, this means the gauge is 10% full. To submit this prompt, the user must then move the slider up and down 203 times (90% of 225). But if the user types 125 words, the gauge is 83% full, so the user only has to move the slider 38 times. If the user types at least 150 words, they do not have to move the slider at all.

Moving the slider up and down to pump up the gauge does not have to be done continuously; the user could lift their mouse after a few movements and continue later. However, when they edit their prompt text, the gauge is reset to reflect the width from words typed. Any changes made by pumping or adding more words to the prompt trigger a 1 s 'pulsing' animation to draw the user's attention to the progress being made. When the gauge is full, the progress bar turns green and the submit button becomes enabled for the user to press. Otherwise, the progress bar is blue and the submit button is disabled.

3.2 Time Cost

We also made minor adjustments to Joshi and Vogel's technique, which introduced a *temporal cost* (cf. [19]; Figure 1b). The technique also has a gauge containing a progress bar with a width that changes as the user types to represent a ratio of the *word minimum*. But this time, if the user's prompt length does not meet the word minimum, the user must press and hold the prompt submission button for a certain time. The time is a ratio of a *wait maximum*, which is 60 s, as was done in their original paper. Unlike the motor cost, the user must press and hold in one act; all progress from pressing and holding is reset to zero when the mouse is released. The progress bar expands accordingly every 100 ms. The same 'pulsing' animation

¹Both techniques are shown in detail in the accompanying video figure.

Table 1: Demographic information for Experiment 1, adapted from Masson et al. [30].

Gender		Age		Education		Creative Writing Frequency		Prompt Engineering Familiarity		
Men	20	18-24	1	High School	3	Weekly	3	Extremely	2	
Women	9	25-34	5	Some University (no credit)	4	Monthly	8	Moderately	4	
		35-44	16	Technical Degree	2	Less than Monthly	9	Somewhat	5	
		45-54	4	Bachelor's Degree	16	Never	9	Slightly	9	
		55-64	1	Master's Degree	4			Not at All	9	
		65-74	2							

ChatGPT Frequency		Weekly ChatGPT Use		Other AI Services				AI Usage				
Daily	11	0 times		3	OpenAI Models	20	Dall-E	3	Write Emails	10	Get a Definition	19
Weekly	11	1-5 times		8	CoPilot	4	Midjourney	2	Write Papers/Essays	4	Explore a Topic	20
Monthly	5	5-10 times		5	Bing	16	Stable Diffusion	1	Write Stories	9	Brainstorming	20
Less than Monthly	2	10-30 times		6	Gemini	19	Other	4	Edit Text	12	Find References	6
		30+ times		7	Grammarly	9			Generate Code	2	Clarification	18
								Edit Code	2	Translate Text	9	
								Debug Code	1	Other	1	
								Generate Images	20			

triggers when the user types or is holding the prompt submission button, and the progress bar turns from blue to green when full.

4 Experiment 1: Submission Techniques

The goal of this experiment is to evaluate the effectiveness of the interaction techniques that introduce friction through costs for increasing prompt length and psychological ownership. This is a within-subjects experiment where participants wrote prompts to generate short stories using each technique.

4.1 Participants

We recruited 35 participants from the online crowdsourcing platform, Prolific.² Participants were restricted to the United States and Canada, and those who had successfully completed at least 2,500 tasks on the platform with a 99-100% approval rating. We manually examined all open-ended responses to identify participants who experienced technical issues or who cheated by not making enough keystrokes in the prompt, suggesting they copied and pasted text into the prompt entry text box. Six (17%) were removed from the analysis for these reasons, leaving 29 valid responses (Table 1). All self-reported English reading and writing proficiency (all ≥ 4 on a 7-point scale). Participants received \$7.50 after completing the experiment.

4.2 Apparatus and Task

The experiment interface was served as a Node.js and React web application. It resembled ChatGPT, with a few minor changes (Figure 2). The top of the interface displayed instructions and a “triad” of nouns to write about, which was adapted from Foley et al. [15] and previously used in Joshi and Vogel’s experiment [19]. Writing short stories based on a series of nouns is useful for this type of experiment as it encourages impromptu creative thinking during experiments [13], allows for increased experimental control, and also encourages high personal investment to write. At the bottom, was the prompt entry interface. The prompt text box was 46 pixels

tall by default and grew by 46 pixels as new lines were typed, until a maximum height of 300 pixels. Beside it was the prompt submission button, which was modified, depending on the condition. Underneath the text box was a message that explained how the current technique worked. After writing their prompt and submitting it, the generated story appeared gradually by words or phrases on top of the text box, giving participants a chance to read it. The story was generated using the OpenAI GPT-4.5 Preview model. After, the user pressed a *Continue* button at the bottom right corner to conclude the trial.

4.3 Procedure

The task was restricted to laptops and desktops. Through the Prolific system, participants accessed a link to the web application. First, they entered basic demographic information and read instructions. Next, they wrote a prompt to generate a short story, and then answered questions about their experience. They repeated this for all conditions. Finally, they answered questions about their overall thoughts and preferences. The entire experiment lasted roughly 30 minutes.

4.4 Design

This is a within-subjects experiment with one primary independent variable, *CONDITION* (levels: *BASELINE*, *TIME*, *MOTOR*). We suspected possible order effects of *CONDITION*, where participants may write more for *BASELINE* after experiencing *TIME* or *MOTOR* first. As such, to get a more accurate measure of *BASELINE*, participants always prompted with this *CONDITION* first. The order of *TIME* and *MOTOR* was randomly assigned. There was only one trial per *CONDITION*. The primary measures are subjective questions (all 0-100 interval data).³ *Psychological Ownership* represents how much psychological ownership participants felt towards their stories. Like Joshi and Vogel’s experiment [19], this was done by asking four questions [5, 32] about *Personal Ownership*, *Responsibility*, *Personal Connection*, and *Emotional Connection* and averaging these results to create

²<https://www.prolific.com>

³See the supplementary materials for question wording and experimental data.

Table 2: Omnibus and post hoc statistical tests for Experiment 1: (a) Word Count, (b) Text Similarity, (c) Psychological Ownership, (d) Mental Demand, (e) Physical Demand, (f) Effort, and (g) Frustration.

(a) Word Count				(b) Text Similarity				(c) Psychological Ownership				(d) Mental Demand			
$\chi^2_{2,N=29} = 16.72, p < .001, W = .29$				$\chi^2_{2,N=29} = 7.52, p < .05, W = .13$				$\chi^2_{2,N=29} = 9.95, p < .01, W = .13$				$\chi^2_{2,N=29} = 15.21, p < .001, W = .26$			
comparisons		p-value		comparisons		p-value		comparisons		p-value		comparisons		p-value	
BASELINE	TIME	< .001	***	BASELINE	TIME	.03	*	BASELINE	TIME	.04	*	BASELINE	TIME	.006	**
BASELINE	MOTOR	< .001	***	BASELINE	TIME	.01	*	BASELINE	MOTOR	.04	*	BASELINE	MOTOR	< .001	***
TIME	MOTOR	.37	n.s.	TIME	MOTOR	.67	n.s.	TIME	MOTOR	.47	n.s.	TIME	MOTOR	.92	n.s.

(e) Physical Demand				(f) Effort				(g) Frustration			
$\chi^2_{2,N=29} = 37.90, p < .001, W = .65$				$\chi^2_{2,N=29} = 23.79, p < .001, W = .41$				$\chi^2_{2,N=29} = 16.53, p < .001, W = .28$			
comparisons		p-value		comparisons		p-value		comparisons		p-value	
BASELINE	TIME	< .001	***	BASELINE	TIME	< .001	***	BASELINE	TIME	< .001	***
BASELINE	MOTOR	< .001	***	BASELINE	TIME	< .001	***	BASELINE	MOTOR	.01	*
TIME	MOTOR	.02	*	TIME	MOTOR	.19	n.s.	TIME	MOTOR	.57	n.s.

a composite measure. The internal consistency reliability score was high ($\alpha = .94$), so this composite measure was appropriate to use for analysis. We also asked about workload-related factors from the NASA-TLX: *Mental Demand*, *Physical Demand*, *Temporal Demand*, *Performance*, *Effort*, and *Frustration*. In addition, the following objective measures were collected from logs: *Word Count*, how long the prompt was when it was submitted; *Text Similarity*, the semantic text similarity between the final prompt and generated story, calculated using Google’s Universal Sentence Encoder (0-1 range; 1 for identical texts) [6]; and the *Initial Word Count*, how long the prompt was when the participant first tried to press the button (TIME) or move the slider (MOTOR).

4.5 Results

In the analysis to follow, we use a Friedman test and Wilcoxon signed-rank tests with Holm’s corrections for multiple comparisons, and Spearman’s correlation coefficients. *Detailed test results are shown in Table 2.* All open-ended responses were grouped into broad themes by the first author.

4.5.1 Word Count. Overall, both interaction techniques led to significantly longer prompts (Figure 3). There was a significant effect of CONDITION on *Word Count*, with post hoc tests revealing that TIME (MDN = 35, IQR = 80) and MOTOR (MDN = 49, IQR = 91) both led to longer prompts than BASELINE (MDN = 15, IQR = 23). *Word Count* more than doubled for TIME and tripled for MOTOR. Seven participants (24%) noted how the interaction techniques “gave some

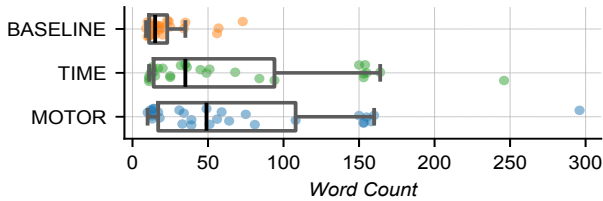


Figure 3: Word Count by CONDITION.

time to reflect” (P14), and encouraged them to “write more and add more detail” (P22).

4.5.2 Text Similarity. Both interaction techniques resulted in higher similarities between the prompt and output. A significant effect of CONDITION on *Text Similarity* and post hoc tests revealed that the generated stories for TIME (MDN = .37, IQR = .18) and MOTOR (MDN = .36, IQR = .12) shared more similarities with the prompts than BASELINE (MDN = .28, IQR = .15). *Text Similarity* was moderately positively correlated with *Word Count* ($r_s = .48, p < .0001$), suggesting that as participants wrote longer prompts, the resulting stories more closely resembled them.

4.5.3 Mechanics. Although there were significant improvements in *Word Count* overall, TIME and MOTOR did not work for all participants. For TIME, participants tried to push the button 10 times on average. When comparing participants’ *Initial Word Count* to their final *Word Count*, participants wrote significantly fewer words initially ($p < .01$; MDN = 20, IQR = 36). Twelve (41%) increased their *Word Count* (MDN = 42.5, IQR = 81.5 for differences), with 4 (14%) increasing by 50 words or more. However, majority (16, 55%) showed no differences between their *Initial Word Count* and final *Word Count*, and of these, only 1 wrote a prompt that was at least 150 words. These results are similar to Joshi and Vogel’s original experiment [19].

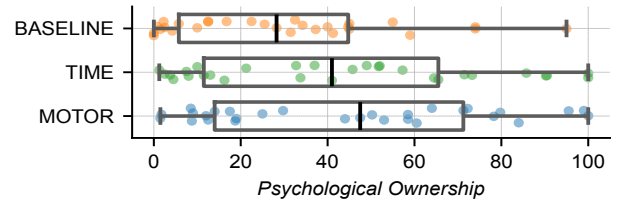


Figure 4: Psychological Ownership by CONDITION. Higher scores correspond to higher feelings of psychological ownership.

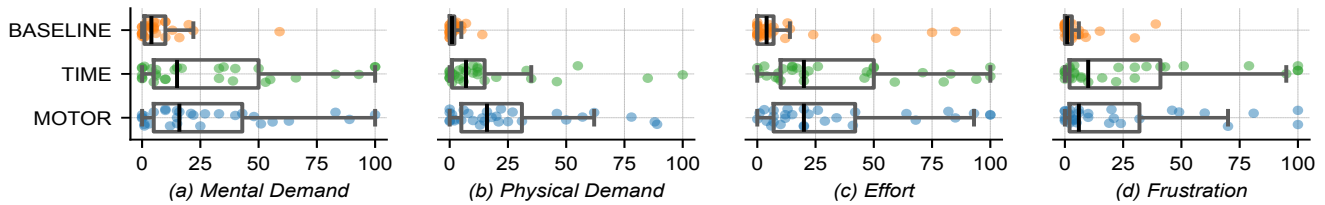


Figure 5: Mental Demand, Physical Demand, Frustration, and Effort by CONDITION.

We observed a similar trend with MOTOR. Participants tried to move the slider 6 times on average. Participants wrote significantly fewer words initially ($p < .01$; MDN = 31, IQR = 28). Twelve (41%) increased their *Word Count* from their *Initial Word Count* (MDN = 31.5, IQR = 52.25 for differences), with 5 (17%) increasing by 50 words or more. Majority of participants (16, 55%) had no difference, and of these, 5 (17%) already wrote prompts that were at least 150 words long, meaning they did not have to move the slider at all.

4.5.4 Psychological Ownership. Both interaction techniques led to significantly higher feelings of psychological ownership (Figure 4). A significant effect of CONDITION on *Psychological Ownership* with post hoc tests revealed that TIME (MDN = 41, IQR = 54) and MOTOR (MDN = 47.5, IQR = 57.25) led to significantly higher *Psychological Ownership* than BASELINE (MDN = 28.25, IQR = 39), which was supported by comments like: “I feel [like the time delay] was prompting me to be more involved in the story details. I would spend more time on the prompt itself. It made the story a bit more personal” (P15).

Although participants wrote longer prompts with TIME and MOTOR, *Word Count* and *Text Similarity* were only weakly positively correlated with *Psychological Ownership* (both $r_s = .28, p < .01$). This may suggest that workload-related factors of the techniques contributed more to the higher feelings of psychological ownership than just prompt length and similarity.

4.5.5 Task Workload. Both interaction techniques led to significantly higher *Mental Demand*, *Physical Demand*, *Effort*, and *Frustration* (Figure 5). For *Mental Demand*, scores were lower for BASELINE (MDN = 4, IQR = 9) than TIME (MDN = 15, IQR = 45) and MOTOR (MDN = 16, IQR = 38). *Mental Demand* was moderately positively correlated with *Word Count* ($r_s = .48, p < .0001$) and *Psychological Ownership* ($r_s = .5, p < .0001$). MOTOR had the highest *Physical Demand* (MDN = 16, IQR = 26), followed by TIME (MDN = 7, IQR = 14), and BASELINE (MDN = 1, IQR = 2). For *Effort*, BASELINE (MDN = 4, IQR = 7) resulted in lower scores than TIME (MDN = 20, IQR = 40) and MOTOR (MDN = 20, IQR = 35). *Effort* was weakly positively correlated with *Word Count* ($r_s = .35, p < .001$) and moderately positively correlated with *Psychological Ownership* ($r_s = .42, p < .0001$). BASELINE (MDN = 1, IQR = 3) had lower *Frustration* than TIME (MDN = 10, IQR = 39) and MOTOR (MDN = 6, IQR = 30). No other metrics or correlations were significant.

These findings suggest that the increased *Mental Demand* and *Effort* required to submit prompts with TIME and MOTOR contributed to higher feelings of *Psychological Ownership*. This was supported by comments like: “I think [the slider] encourages me to add more of a personal input in the story that is submitted. Although it was a bit annoying, the end result felt more of a personal effort than not” (P20);

and “having to move the slider handle up and down felt like work because it’s something I had to do in order to submit the prompt. It made me feel a slightly deeper connection since I sort of had to work for it” (P18).

4.6 Summary

Overall, our results show that increasing the temporal and motor cost of submitting a prompt can encourage users to write longer prompts. These techniques required additional mental and physical demand, and were more effortful and frustrating. However, the increased prompt length, text similarity, mental demand, and effort of these techniques all contributed to higher feelings of psychological ownership. These findings validate and significantly extend those of Joshi and Vogel’s initial experiment [19] and provide evidence that incorporating motor costs into the prompt submission process can improve psychological ownership.

5 Receiving Suggestions from AI

Encouraged by these results, we were interested in further augmenting these techniques to see if they could further improve psychological ownership. Recall that prior work shows that receiving suggestions from AI can lower [10, 26] or increase [27] psychological ownership depending on the interaction. Even in writing scenarios that do not involve AI, receiving feedback can help writers [23] and suggestions can encourage ideation and creativity, and help to overcome writer’s block [36], which may occur when writing longer prompts [19]. Receiving suggestions while waiting or moving the slider may further encourage users to write longer prompts, which may further increase mental demand, effort, and psychological ownership.

We propose augmenting these techniques to display suggestions for how the prompt can be expanded when the user stops writing. After the user stops writing for 1 second, a blue tooltip appears on the top right corner of the prompt submission box (Figure 1c). This resembles the “on-idle” condition from Siangliulue et al. [36]. The tooltip contains a suggestion, phrased as a question to encourage more thought [9], that is one sentence long for how the prompt can be expanded or improved to incorporate more details (e.g., *Consider: how does the relationship between the characters influence the story?*) These suggestions are generated using GPT 3.5 Turbo for increased speed.⁴ Importantly, the suggestion is not automatically incorporated as edits to the prompt as this may reduce psychological ownership [27]. The tooltip stays on the screen at maximum opacity for 4 seconds, then gradually fades away for 3 seconds, mimicking the mechanics of prior work on displaying notifications [18]. This

⁴See the supplementary materials for the prompt used to generate the suggestion.

Table 3: Demographic information for Experiment 2, adapted from Masson et al. [30].

Gender		Age		Education		Creative Writing Frequency		Prompt Engineering Familiarity	
Men	15	18-24	2	High School	4	Daily	1	Extremely	3
Women	15	25-34	6	Some University (no credit)	2	Weekly	6	Moderately	6
		35-44	6	Technical Degree	2	Monthly	6	Somewhat	5
		45-54	10	Bachelor's Degree	17	Less than Monthly	12	Slightly	7
		55-64	4	Master's Degree	4	Never	5	Not at All	9
		65-74	2	Doctorate	1				

ChatGPT Frequency		Weekly ChatGPT Use		Other AI Services		AI Usage			
Daily	8	0 times	4	Nothing	2	Grammarly	12	Nothing	2
Weekly	12	1-5 times	12	OpenAI Models	23	Dall-E	5	Write Emails	14
Monthly	5	5-10 times	5	CoPilot	4	Midjourney	4	Write Papers/Essays	7
Less than Monthly	3	10-30 times	6	Bing	12	Stable Diffusion	3	Write Stories	9
Never	2	30+ times	3	Gemini	18	Other	3	Edit Text	17
								Generate Code	5
								Edit Code	4
								Debug Code	3
								Generate Images	15
								Get a Definition	17
								Explore a Topic	17
								Brainstorming	20
								Find References	9
								Clarification	13
								Translate Text	9
								Other	1

gives users time to choose whether they would like to incorporate these suggestions or not. If the user begins typing again, the tooltip does not show up (unless they again stop typing for 1 second). Otherwise, the tooltip reappears after 1 second.

6 Experiment 2: Prompt Suggestions

We recruited 30 new participants using the same inclusion and exclusion criteria as before. No participants had to be filtered out (Table 3). Participants on Prolific are known for producing high-quality data [11], which could explain why no data had to be dropped. All self-reported English reading and writing proficiency (all ≥ 4 on a 7-point scale). The task and procedure were both identical to that of Experiment 1. The apparatus was mostly the same; however, each technique was further enhanced to display suggestions (Figure 1c). These suggestions were integrated into the same interaction techniques and baseline as before. For all conditions, the prompt submission button was disabled until at least one suggestion was shown. The design was mostly identical to that of Experiment 1; however, we also considered *Number of Suggestions*, the number of suggestions each participant saw for a single CONDITION.

6.1 Results

We use the same statistical tests as Experiment 1, with *detailed statistical test results in Table 4*.

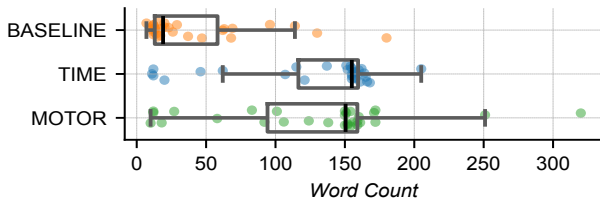


Figure 6: Word Count by CONDITION when receiving AI-generated suggestions.

6.1.1 Word Count. Both interaction techniques led to significantly longer prompts (Figure 6). A significant effect of CONDITION on *Word Count* and post hoc tests showed that BASELINE (MDN = 19, IQR = 45.25) resulted in shorter prompts than TIME (MDN = 155, IQR = 43) and MOTOR (MDN = 150.5, IQR = 64.75). This represents a difference of over 100 words.

6.1.2 Text Similarity. Both interaction techniques also led to more text similarity. A significant effect of *Text Similarity* on CONDITION revealed that prompts written with BASELINE (MDN = .32, IQR = .14) were less similar than TIME (MDN = .41, IQR = .13) and MOTOR (MDN = .42, IQR = .14). *Text Similarity* was moderately positively correlated with *Word Count* ($r_s = .55, p < .0001$).

6.1.3 Mechanics. For TIME, participants pressed the button 7 times on average. When comparing the *Initial Word Count* to the final *Word Count*, participants wrote significantly fewer words initially ($p < .001$; MDN = 91.5, IQR = 126.5). Eleven (37%) were consistent, of which 7 (23%) consistently wrote at least 150 words. Majority of participants (18, 60%) wrote more (MDN = 63, IQR = 110), with 11 (37%) writing 50 words or more.

For MOTOR, participants moved the slider 6 times on average. Again, participants wrote significantly fewer words initially ($p < .01$; MDN = 96.5, IQR = 134.75). Majority (20, 67%) were consistent, with 13 (43%) writing at least 150 words and not having to interact with the slider at all. Ten (33%) wrote more (MDN = 85, IQR = 53.5 for differences), with 8 (27%) writing at least 50 words.

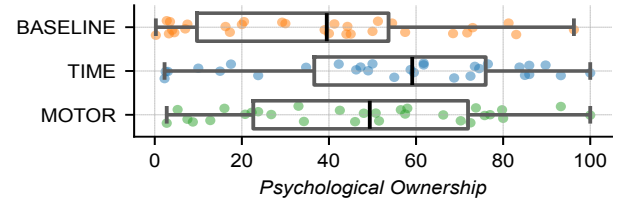
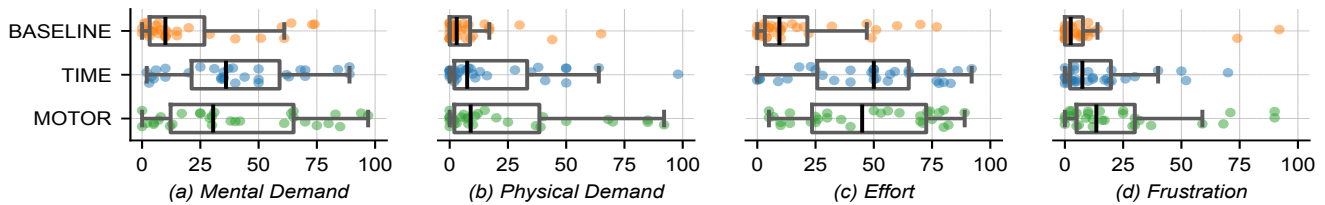


Figure 7: Psychological Ownership by CONDITION when receiving AI-generated suggestions.

Table 4: Omnibus and post hoc statistical tests for Experiment 2: (a) Word Count, (b) Text Similarity, (c) Psychological Ownership, (d) Mental Demand, (e) Physical Demand, (f) Effort, (g) Frustration, and (h) Number of Suggestions.

(a) Word Count				(b) Text Similarity				(c) Psychological Ownership				(d) Mental Demand			
$\chi^2_{2,N=30} = 35.06, p < .001, W = .58$				$\chi^2_{2,N=30} = 14.6, p < .001, W = .24$				$\chi^2_{2,N=30} = 14.36, p < .001, W = .24$				$\chi^2_{2,N=30} = 10.14, p < .01, W = .17$			
comparisons		p-value		comparisons		p-value		comparisons		p-value		comparisons		p-value	
BASELINE	TIME	< .001	***	BASELINE	TIME	< .001	***	BASELINE	TIME	< .001	***	BASELINE	TIME	.003	**
BASELINE	MOTOR	< .001	***	BASELINE	TIME	.002	**	BASELINE	MOTOR	.005	**	BASELINE	MOTOR	.007	**
TIME	MOTOR	.83	n.s.	TIME	MOTOR	.79	n.s.	TIME	MOTOR	.02	*	TIME	MOTOR	.98	n.s.
(e) Physical Demand				(f) Effort				(g) Frustration				(h) Number of Suggestions			
$\chi^2_{2,N=30} = 10.82, p < .01, W = .18$				$\chi^2_{2,N=30} = 23.54, p < .001, W = .39$				$\chi^2_{2,N=30} = 8.90, p < .05, W = .15$				$\chi^2_{2,N=30} = 41.30, p < .001, W = .69$			
comparisons		p-value		comparisons		p-value		comparisons		p-value		comparisons		p-value	
BASELINE	TIME	.04	*	BASELINE	TIME	< .001	***	BASELINE	TIME	.05	n.s.	BASELINE	TIME	< .001	***
BASELINE	MOTOR	.007	**	BASELINE	TIME	< .001	***	BASELINE	MOTOR	.003	**	BASELINE	MOTOR	< .001	***
TIME	MOTOR	.14	n.s.	TIME	MOTOR	.74	n.s.	TIME	MOTOR	.05	n.s.	TIME	MOTOR	.85	n.s.

**Figure 8: Mental Demand, Physical Demand, Frustration, and Effort by CONDITION when receiving AI-generated suggestions.**

6.1.4 Psychological Ownership. Both interaction techniques led to higher feelings of psychological ownership (Figure 7). A significant effect of CONDITION on *Psychological Ownership* and post hoc tests showed that TIME (MDN = 59.12, IQR = 39.38) led to the highest *Psychological Ownership*, followed by MOTOR (MDN = 49.38, IQR = 49.31), and BASELINE (MDN = 39.5, IQR = 44). *Psychological Ownership* was weakly positively correlated with *Word Count* ($r_s = .37, p < .001$) and *Text Similarity* ($r_s = .32, p < .01$).

6.1.5 Task Workload. Both interaction techniques increased *Mental Demand*, *Physical Demand*, *Effort*, and *Frustration* (Figure 8). For *Mental Demand*, BASELINE (MDN = 10, IQR = 23.5) resulted in lower scores than both TIME (MDN = 36, IQR = 37.75) and MOTOR (MDN = 30.5, IQR = 52.75). *Mental Demand* was moderately positively correlated with *Word Count* and strongly positively correlated with *Psychological Ownership* ($r_s = .62, p < .0001$). BASELINE (MDN = 3, IQR = 8.75) required less *Physical Demand* than TIME (MDN = 7.5, IQR = 31.25) and MOTOR (MDN = 9, IQR = 36.5). There was a weak positive correlation between *Physical Demand* and *Psychological Ownership* ($r_s = .35, p < .001$). For *Effort*, BASELINE (MDN = 9.5, IQR = 18.25) was less demanding than TIME (MDN = 50, IQR = 39.25) and MOTOR (MDN = 45, IQR = 49). *Effort* was moderately positively correlated with *Word Count* ($r_s = .43, p < .0001$) and *Psychological Ownership* ($r_s = .5, p < .0001$). MOTOR (MDN = 13.5, IQR = 25) resulted in higher *Frustration* than BASELINE (MDN = 2.5, IQR = 7.75). No other metrics or correlations were statistically significant.

6.1.6 Number of Suggestions. Participants received more suggestions when it was integrated into the two interaction techniques. A

significant effect of CONDITION on *Number of Suggestions* and post hoc tests revealed that participants received fewer suggestions with BASELINE (MDN = 4.5, IQR = 8.25) than TIME (MDN = 22, IQR = 12) and MOTOR (MDN = 23, IQR = 14.5). *Number of Suggestions* was moderately positively correlated with *Word Count* ($r_s = .55, p < .0001$), *Text Similarity* ($r_s = .45, p < .0001$), *Mental Demand* ($r_s = .49, p < .0001$), and *Effort* ($r_s = .54, p < .0001$); and weakly positively correlated with *Physical Demand* ($r_s = .22, p < .05$), *Frustration* ($r_s = .25, p < .05$), and *Psychological Ownership* ($r_s = .26, p < .05$).

Although four participants (13%) noted that the suggestions were “distracting or annoying” (P17), 24 (80%) appreciated them, as they helped when participants “started to feel a bit stuck” (P10) and encouraged more thought, for example: “[the suggestions] got me really thinking and making this more of a collaborative process than just entering a prompt and taking what the AI generated” (P28).

6.2 The Effects of Receiving Suggestions

We use Mann-Whitney U tests to make between-subjects comparisons with Experiment 1 to examine the effect of receiving AI-generated suggestions (levels: NONE, from Experiment 1; and SUGGESTIONS, from Experiment 2). *Statistical test details are in Table 5*.

Overall, participants wrote over 100 more words with both interaction techniques when they received suggestions. For both TIME and MOTOR, *Word Count* was significantly higher with SUGGESTIONS, but receiving suggestions did not affect BASELINE. We observed a similar trend for *Text Similarity*, where SUGGESTIONS increased

Table 5: Between-subjects statistical tests comparing NONE and SUGGESTIONS for: (a) BASELINE, (b) TIME, and (c) MOTOR.

	(a) BASELINE		(b) TIME		(c) MOTOR	
<i>metric</i>	<i>p-value</i>		<i>p-value</i>		<i>p-value</i>	
<i>Word Count</i>	.10	<i>n.s.</i>	< .001	***	.003	**
<i>Text Similarity</i>	.22	<i>n.s.</i>	.04	*	.02	*
<i>Initial Word Count</i>			.01	*	.44	<i>n.s.</i>
<i>Psychological Ownership</i>	.25	<i>n.s.</i>	.16	<i>n.s.</i>	.73	<i>n.s.</i>
<i>Mental Demand</i>	.008	**	.10	<i>n.s.</i>	.11	<i>n.s.</i>
<i>Physical Demand</i>	.03	*	.92	<i>n.s.</i>	.46	<i>n.s.</i>
<i>Temporal Demand</i>	.01	*	.22	<i>n.s.</i>	.20	<i>n.s.</i>
<i>Effort</i>	.04	*	.10	<i>n.s.</i>	.07	<i>n.s.</i>

similarity for TIME and MOTOR, but not BASELINE. For TIME, participants made their first submission attempt when *Initial Word Count* was significantly longer with SUGGESTIONS, but this did not occur for MOTOR. This could explain why there were significant differences in *Psychological Ownership* between TIME and MOTOR for SUGGESTIONS.

Even though participants wrote significantly more words and had more similar texts with TIME and MOTOR, this did not further improve *Psychological Ownership* as no comparisons with NONE were significant. For BASELINE, having SUGGESTIONS led to significantly higher *Mental Demand*, *Physical Demand*, *Temporal Demand*, and *Effort*, but there were no differences for TIME and MOTOR.

6.3 Summary

Overall, our results show that providing AI-generated suggestions can help users write even longer prompts when using techniques that increase the temporal and motor costs of submitting prompts. However, this did not further improve psychological ownership when compared to providing no suggestions.

7 Discussion

Our first experiment shows that introducing friction by increasing the time and motor costs of submitting short prompts can encourage users to write longer prompts. This typically required more mental demand and effort, but critically, it also increased psychological ownership. Our second experiment further enhanced these two techniques by incorporating AI-generated suggestions into the prompt entry interface. This further increased prompt length but did not further improve psychological ownership when compared to no suggestions.

These findings expand and provide additional nuance to Joshi and Vogel’s original experiment [19]. Even though receiving suggestions made it easier for participants to write more and resulted in more text similarity, this did not lead to differences in psychological ownership. One possibility is that psychological ownership plateaued with longer prompts [19], but another is that other factors beyond prompt length and text similarity were critical to increasing it. It could have been that workload-related factors were the most important contributors to increased psychological ownership. Unlike word count and text similarity, which were only weakly positively correlated with psychological ownership, mental demand and effort were notably both moderately positively correlated with psychological ownership in both experiments. To verify this, we

conducted a variable importance analysis using the random forest method, and found that mental demand was the most important contributor to increased psychological ownership. Prior work also suggests that receiving suggestions from AI can affect what users think and write about [17], so another possibility is that receiving suggestions affected how participants thought about their stories and felt less psychological ownership as a result.

This is not to say that encouraging users to write prompt length is not a valuable outcome; rather, encouraging users to write longer prompts can be a way to encourage more thought and improve psychological ownership. As such, future work could continue to explore ways to increase prompt length and mental demand by imposing other costs when submitting short prompts, like perceptual costs, which have shown promise in non-AI contexts [7, 38].

We believe these ideas of increasing the costs of submitting shorter prompts could be valuable for other types of human-AI collaborations, like requiring users to submit detailed sketches when generating images to improve higher psychological ownership. It could also be useful in non-AI writing contexts. For example, incorporating such techniques into survey responses could encourage respondents to write longer open-ended responses.

Encouraging users to write longer prompts is just one technique for increasing psychological ownership that can be easily integrated into current, chat-based interfaces. Future work should explore other mechanisms that encourage users to put more thought and effort into their interactions with generative AI writing assistants for increased psychological ownership.

7.1 Limitations

The baseline condition was always presented first to reduce potential order effects (writing more in BASELINE after experiencing TIME or MOTOR first). However, we acknowledge that always experiencing the baseline condition first may have primed participants in ways that could have influenced their performance in subsequent conditions (e.g., by increasing task familiarity or becoming more inspired to write after completing the baseline trial).

Our results from the second experiment may also be limited by how we chose to display the prompt suggestions. Given the short writing task and the design of the two interaction techniques we tested, we felt that showing suggestions after 1 s of inactivity was appropriate; however, it may have been more appropriate to display them after a longer period or to display them on-demand (e.g., after the participant presses a button [36, 37]). On-demand suggestions

may also result in increased psychological ownership due to the explicit action required from the user to receive a suggestion [27].

An ongoing challenge is integrating such techniques into generative AI writing tools. Users may not be willing engage with them for a real-world writing task that requires more urgency, but prior work suggests that additional interface augmentations could help, for example, by providing feedback and explanations for why these techniques exist [14, 31]. In practice, these techniques may be better suited to being invoked at the user's request, aligning with ongoing trends to self-impose 'lockout' mechanisms to manage technology use (e.g., [21, 22]).

8 Conclusion

We propose increasing the motor cost of submitting shorter prompts by requiring users to repeatedly move a slider up and down to fill a 'gauge,' which can also be filled by writing more words in the prompt submission text box. Two experiments compared this approach to a technique from prior work that increases the temporal cost by requiring users to press and hold the prompt submission button and a baseline technique with no augmentations, with and without AI-generated suggestions. Both types of costs introduced friction to encourage users to write longer prompts and increased psychological ownership, but despite the even longer prompts written with AI-generated suggestions, this did not further improve psychological ownership. Our results show that simple interaction techniques that introduce friction to increase the time and motor costs of submitting short prompts can encourage users to submit longer prompts with more details, which can increase mental demand and positively impact psychological ownership and human-AI collaborative writing.

References

- [1] Oloff C. Biemann, Ning F. Ma, and Dongwook Yoon. 2022. From Tool to Companion: Storywriters Want AI Writers to Respect Their Personal Values and Writing Strategies. In *Proceedings of the 2022 ACM Designing Interactive Systems Conference* (Virtual Event, Australia) (DIS '22). Association for Computing Machinery, New York, NY, USA, 1209–1227. doi:10.1145/3532106.3533506
- [2] Ina Blau and Avner Caspi. 2009. What Type of Collaboration Helps? Psychological Ownership, Perceived Learning and Outcome Quality of Collaboration using Google Docs. In *Proceedings of the Chais Conference on Instructional Technologies Research*, Vol. 12. 48–55.
- [3] Zana Bucinca, Maja Barbara Malaya, and Krzysztof Z. Gajos. 2021. To Trust or to Think: Cognitive Forcing Functions Can Reduce Overreliance on AI in AI-assisted Decision-making. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1, Article 188 (April 2021), 21 pages. doi:10.1145/3449287
- [4] Ana Caraban, Evangelos Karapanos, Daniel Gonçalves, and Pedro Campos. 2019. 23 Ways to Nudge: A Review of Technology-Mediated Nudging in Human-Computer Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) (CHI '19). Association for Computing Machinery, New York, NY, USA, 1–15. doi:10.1145/3290605.3300733
- [5] Avner Caspi and Ina Blau. 2011. Collaboration and Psychological Ownership: How does the Tension between the Two Influence Perceived Learning? *Social Psychology of Education* 14 (2011), 283–298.
- [6] Daniel Cer, Yinfei Yang, Sheng-yi Kong, Nan Hua, Nicole Limtiaco, Rhomni St. John, Noah Constant, Mario Guajardo-Cespedes, Steve Yuan, Chris Tar, Brian Strope, and Ray Kurzweil. 2018. Universal Sentence Encoder for English. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, Eduardo Blanco and Wei Lu (Eds.). Association for Computational Linguistics, Brussels, Belgium, 169–174. doi:10.18653/v1/D18-2029
- [7] Andy Cockburn, Per Ola Kristensson, Jason Alexander, and Shumin Zhai. 2007. Hard Lessons: Effort-Inducing Interfaces Benefit Spatial Learning. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (CHI '07). Association for Computing Machinery, New York, NY, USA, 1571–1580. doi:10.1145/1240624.1240863
- [8] Anna L. Cox, Sandy J.J. Gould, Marta E. Cecchinato, Ioanna Iacovides, and Ian Renfree. 2016. Design Frictions for Mindful Interactions: The Case for Microboundaries. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (San Jose, California, USA) (CHI EA '16). Association for Computing Machinery, New York, NY, USA, 1389–1397. doi:10.1145/2851581.2892410
- [9] Valdemar Danry, Pat Pataranutaporn, Yaoli Mao, and Pattie Maes. 2023. Don't Just Tell Me, Ask Me: AI Systems that Intelligently Frame Explanations as Questions Improve Human Logical Discernment Accuracy over Causal AI explanations. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) (CHI '23). Association for Computing Machinery, New York, NY, USA, Article 352, 13 pages. doi:10.1145/3544548.3580672
- [10] Paramveer S. Dhillon, Somayeh Molaei, Jiaqi Li, Maximilian Golub, Shaochun Zheng, and Lionel Peter Robert. 2024. Shaping Human-AI Collaboration: Varied Scaffolding Levels in Co-writing with Language Models. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '24). Association for Computing Machinery, New York, NY, USA, Article 1044, 18 pages. doi:10.1145/3613904.3642134
- [11] Benjamin D. Douglas, Patrick J. Ewell, and Markus Brauer. 2023. Data quality in online human-subjects research: Comparisons between MTurk, Prolific, CloudResearch, Qualtrics, and SONA. *PLOS ONE* 18, 3 (03 2023), 1–17. doi:10.1371/journal.pone.0279720
- [12] Fiona Draxler, Anna Werner, Florian Lehmann, Matthias Hoppe, Albrecht Schmidt, Daniel Buschek, and Robin Welsch. 2024. The AI Ghostwriter Effect: When Users do not Perceive Ownership of AI-Generated Text but Self-Declare as Authors. *ACM Trans. Comput.-Hum. Interact.* 31, 2, Article 25 (feb 2024), 40 pages. doi:10.1145/3637875
- [13] Mark D. Dunlop, Emma Nicol, Andreas Komninos, Prima Dona, and Naveen Durga. 2016. Measuring Inviscid Text Entry using Image Description Tasks. In *Proceedings of the 2016 CHI Workshop on Inviscid Text Entry and Beyond*.
- [14] Serge Egelman, David Molnar, Nicolas Christin, Alessandro Acquisti, Cormac Herley, and Shriram Krishnamurthi. 2010. Please Continue to Hold. In *Ninth Workshop on the Economics of Information Security*.
- [15] Margaret Foley, Géry Casiez, and Daniel Vogel. 2020. Comparing Smartphone Speech Recognition and Touchscreen Typing for Composition and Transcription. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–11. doi:10.1145/3313831.3376861
- [16] Aaron Halfaker, Aniket Kittur, Robert Kraut, and John Riedl. 2009. A Jury of your Peers: Quality, Experience and Ownership in Wikipedia. In *Proceedings of the 5th International Symposium on Wikis and Open Collaboration* (Orlando, Florida) (WikiSym '09). Association for Computing Machinery, New York, NY, USA, Article 15, 10 pages. doi:10.1145/1641309.1641332
- [17] Maurice Jakesch, Advait Bhat, Daniel Buschek, Lior Zalmanson, and Mor Naaman. 2023. Co-Writing with Opinionated Language Models Affects Users' Views. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) (CHI '23). Association for Computing Machinery, New York, NY, USA, Article 111, 15 pages. doi:10.1145/3544548.3581196
- [18] Nikhita Joshi, Richard Li, Jiannan Li, Leonardo Pavanatto, Michel Pahud, Jatin Sharma, Bongshin Lee, Hugo Romat, William Buxton, Nicolai Marquardt, Ken Hinckley, and Nathalie Henry Riche. 2024. Opportunistic Nudges for Task Migration Between Personal Devices. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI EA '24). Association for Computing Machinery, New York, NY, USA, Article 256, 8 pages. doi:10.1145/3613905.3651037
- [19] Nikhita Joshi and Daniel Vogel. 2025. Writing with AI Lowers Psychological Ownership, but Longer Prompts Can Help. In *Proceedings of the 7th ACM Conference on Conversational User Interfaces* (Waterloo, ON, Canada) (CUI '25). Association for Computing Machinery, New York, NY, USA, 17 pages. doi:10.1145/3719160.3736608
- [20] Daniel Kahneman. 2011. *Thinking, Fast and Slow*. Farrar, Straus and Giroux, New York.
- [21] Jaejeung Kim, Joonyoung Park, Hyunsoo Lee, Minsam Ko, and Uichin Lee. 2019. LocknType: Lockout Task Intervention for Discouraging Smartphone App Use. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) (CHI '19). Association for Computing Machinery, New York, NY, USA, 1–12. doi:10.1145/3290605.3300927
- [22] Minsam Ko, Seungwoo Choi, Koji Yatani, and Uichin Lee. 2016. Lock n' LoL: Group-based Limiting Assistance App to Mitigate Smartphone Distractions in Group Activities. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (CHI '16). Association for Computing Machinery, New York, NY, USA, 998–1010. doi:10.1145/2858036.2858568
- [23] Sneha R. Krishna Kumaran, Yue Yin, and Brian P. Bailey. 2021. Plan Early, Revise More: Effects of Goal Setting and Perceived Role of the Feedback Provider on Feedback Seeking Behavior. 5, CSCW1, Article 24 (April 2021), 22 pages. doi:10.1145/3449098

- [24] Ida Larsen-Ledet and Henrik Korsgaard. 2019. Territorial Functioning in Collaborative Writing: Fragmented Exchanges and Common Outcomes. *Comput. Supported Coop. Work* 28, 3–4 (Jun 2019), 391–433. doi:10.1007/s10606-019-09359-8
- [25] Mina Lee, Katy Ilonka Gero, John Joon Young Chung, Simon Buckingham Shum, Vipul Raheja, Hua Shen, Subhashini Venugopalan, Thimo Wambsganss, David Zhou, Emad A. Alghamdi, Tal August, Avinash Bhat, Madiha Zahrah Choksi, Senjuti Dutta, Jin L.C. Guo, Md Naimul Hoque, Yewon Kim, Simon Knight, Seyed Parsa Neshaei, Antonette Shibani, Disha Shrivastava, Lila Shroff, Agnia Sergeyuk, Jessi Stark, Sarah Sterman, Sitong Wang, Antoine Bosselut, Daniel Buschek, Joseph Chee Chang, Sherol Chen, Max Kreminski, Joonsuk Park, Roy Pea, Eugenia Ha Rim Rho, Zejiang Shen, and Pao Siangliulue. 2024. A Design Space for Intelligent and Interactive Writing Assistants. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '24). Association for Computing Machinery, New York, NY, USA, Article 1054, 35 pages. doi:10.1145/3613904.3642697
- [26] Mina Lee, Percy Liang, and Qian Yang. 2022. CoAuthor: Designing a Human-AI Collaborative Writing Dataset for Exploring Language Model Capabilities. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) (CHI '22). Association for Computing Machinery, New York, NY, USA, Article 388, 19 pages. doi:10.1145/3491102.3502030
- [27] Florian Lehmann, Niklas Markert, Hai Dang, and Daniel Buschek. 2022. Suggestion Lists vs. Continuous Generation: Interaction Design for Writing with Generative Models on Mobile Devices Affect Text Length, Wording and Perceived Authorship. In *Proceedings of Mensch Und Computer 2022* (Darmstadt, Germany) (MuC '22). Association for Computing Machinery, New York, NY, USA, 192–208. doi:10.1145/3543758.3543947
- [28] Blaine Lewis and Daniel Vogel. 2020. Longer Delays in Rehearsal-based Interfaces Increase Expert Use. *ACM Trans. Comput.-Hum. Interact.* 27, 6, Article 45 (nov 2020), 41 pages. doi:10.1145/3418196
- [29] Zhuoyan Li, Chen Liang, Jing Peng, and Ming Yin. 2024. The Value, Benefits, and Concerns of Generative AI-Powered Assistance in Writing. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '24). Association for Computing Machinery, New York, NY, USA, Article 1048, 25 pages. doi:10.1145/3613904.3642625
- [30] Damien Masson, Sylvain Malacria, Géry Casiez, and Daniel Vogel. 2024. Direct-GPT: A Direct Manipulation Interface to Interact with Large Language Models. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '24). Association for Computing Machinery, New York, NY, USA, Article 975, 16 pages. doi:10.1145/3613904.3642462
- [31] Fiona Fui-Hoon Nah. 2004. A Study on Tolerable Waiting Time: How Long are Web Users Willing to Wait? *Behaviour & Information Technology* 23, 3 (2004), 153–163.
- [32] Justin Nicholes. 2017. Measuring Ownership of Creative Versus Academic Writing: Implications for Interdisciplinary Praxis. *Writing in Practice* 3. doi:10.62959/WIP-03-2017-08
- [33] Jon L. Pierce, Tatiana Kostova, and Kurt T. Dirks. 2001. Toward a Theory of Psychological Ownership in Organizations. *The Academy of Management Review* 26, 2 (2001), 298–310. doi:10.2307/259124
- [34] Jon L. Pierce, Tatiana Kostova, and Kurt T. Dirks. 2003. The State of Psychological Ownership: Integrating and Extending a Century of Research. *Review of General Psychology* 7, 1 (2003), 84–107. doi:10.1037/1089-2680.7.1.84
- [35] Hua Shen and Tongshuang Wu. 2023. Parachute: Evaluating Interactive Human-LM Co-writing Systems. arXiv:2303.06333
- [36] Pao Siangliulue, Joel Chan, Krzysztof Z. Gajos, and Steven P. Dow. 2015. Providing Timely Examples Improves the Quantity and Quality of Generated Ideas. In *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition* (Glasgow, United Kingdom) (C&C '15). Association for Computing Machinery, New York, NY, USA, 83–92. doi:10.1145/2757226.2757230
- [37] Nikhil Singh, Guillermo Bernal, Daria Savchenko, and Elena L. Glassman. 2023. Where to Hide a Stolen Elephant: Leaps in Creative Writing with Multimodal Machine Intelligence. 30, 5, Article 68 (Sept. 2023), 57 pages. doi:10.1145/3511599
- [38] Frank Soboczinski, Paul Cairns, and Anna L. Cox. 2013. Increasing Accuracy by Decreasing Presentation Quality in Transcription Tasks. In *Human-Computer Interaction – INTERACT 2013* (Cape Town, South Africa). Springer Berlin Heidelberg, Berlin, Heidelberg, 380–394.
- [39] Jennifer Thom-Santelli, Dan R. Cosley, and Geri Gay. 2009. What's Mine is Mine: Territoriality in Collaborative Authoring. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA) (CHI '09). Association for Computing Machinery, New York, NY, USA, 1481–1484. doi:10.1145/1518701.1518925
- [40] QianYing Wang, Alberto Battocchi, Ilenia Graziola, Fabio Pianesi, Daniel Tomasini, Massimo Zancanaro, and Clifford Nass. 2006. The Role of Psychological Ownership and Ownership Markers in Collaborative Working Environment (ICMI '06). Association for Computing Machinery, New York, NY, USA, 225–232. doi:10.1145/1180995.1181041
- [41] Yang Wang, Pedro Giovanni Leon, Alessandro Acquisti, Lorrie Faith Cranor, Alain Forget, and Norman Sadeh. 2014. A Field Trial of Privacy Nudges for Facebook. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*

(Toronto, Ontario, Canada) (CHI '14). Association for Computing Machinery, New York, NY, USA, 2367–2376. doi:10.1145/2556288.2557413

A Pilot Experiment

We wanted to make sure that the two interaction techniques were similar in how long they took, as differences could have confounded our results. As the time cost technique required users to wait for 60 s in the extreme case, we wanted the extreme case of the motor cost technique to take roughly 60 s as well. To get a sense of how many slider movements would be needed to approximate 60 s, we asked 5 people within our research group to try moving the slider 100 times. The median duration for a single movement was 267 ms. Dividing 60,000 ms by 267 ms gives roughly 225 movements to approximate 60 s.