Prompting by Doing: How Direct Manipulation can Protect and Augment Writers' Thoughts in AI Tools

Tim Zindulka
tim.zindulka@uni-bayreuth.de
University of Bayreuth
Bayreuth, Germany

Jannek Sekowski
jannek.sekowski@uni-bayreuth.de
University of Bayreuth
Bayreuth, Germany

Sven Goller
sven.goller@uni-bayreuth.de
University of Bayreuth
Bayreuth, Germany

Florian Lehmann
florian.lehmann@uni-bayreuth.de
University of Bayreuth
Bayreuth, Germany

Daniel Buschek
daniel.buschek@uni-bayreuth.de
University of Bayreuth
Bayreuth, Germany

Abstract

AI tools for text generation shape how people write – and how they think. Current turn-based interactions (e.g. chatbots) force users to write prompts, switch UIs and writing mindsets to do so, and thus add cognitive overhead to the process. In this position paper, we argue that *direct manipulation* offers a powerful alternative approach for integrating AI features in a way that augments and protects people's writing workflows and thinking: Instead of using a dedicated UI for prompting, users manipulate interactive objects embedded in their writing environment, which represent user intents and AI responses. To explore this idea, we reflect on direct manipulation techniques and their impact on cognition in recent related research and our own studies on writing tools. Overall, we highlight direct manipulation as a promising interaction concept for generative AI systems – to reduce prompt writing, ease cognitive demands, and eliminate unnecessary workflow disruptions and context switches.

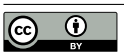
CCS Concepts

• **Human-centered computing** → **Empirical studies in HCI**; **Text input**; • **Computing methodologies** → **Natural language processing**.

Keywords

Writing assistance, Large language models, Human-AI interaction, Mobile interaction, Touch interaction, Direct manipulation, Cognition

ACM Reference Format:
Tim Zindulka, Jannek Sekowski, Sven Goller, Florian Lehmann, and Daniel Buschek. 2025. Prompting by Doing: How Direct Manipulation can Protect and Augment Writers' Thoughts in AI Tools. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26-May 1, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 5 pages. https://doi.org/XX.XXXX/XXXXXXX.XXXXXXX



CHI '25, Yokohama, Japan

ACM ISBN XXX-X-XXXX-XXXX-X/XX/XX
https://doi.org/XX.XXXX/XXXXXXX.XXXXXXX

1 Introduction & Position

Ironically, obtaining writing support from a Large Language Model (LLM) requires more writing. Now, imagine you've finally found a quiet minute to draft on your CHI paper but you keep interrupting your writing to message todos to a colleague. You've created parallel writing tasks that may both be productive but perhaps less so in combination. That's what it feels like when prompting AI writing assistants today.

We take the position that this interaction style, which forces writers to interrupt their writing process and craft prompts, is poorly suited to augment and protect their thought processes. Instead, we argue that direct manipulation (DM), a category of interaction techniques where users can directly manipulate on-screen objects (e.g., through clicking, drawing, or dragging) instead of relying on textual commands, can better support the use of generative AI (GenAI) while thinking and writing. This stems from two main observations:

(1) Both writing and prompting require thinking. As prior research has shown, writing effective prompts is challenging [10], particularly when switching between working on the draft and writing prompts [4]. Putting complex ideas and nuanced intentions into prompts produces cognitive overhead on writers, disrupting the flow of their thoughts. Direct manipulation helps solve this challenge by letting users generate prompts through incremental interaction with a continuously updated representation of the text, instead of typing out instructions. For example, writers could expand a paragraph by stretching/dragging it to their desired length or pinching it to make it more concise. This removes the need to copy and paste text between apps or craft specific prompts like "add only few words here" or "simplify this section but keep all key arguments." In addition, if users write prompts, DM techniques can reify them, which better respects the time and effort users invested in crafting these prompts. For example, writers may turn frequently used prompts into reusable, interactive elements, like buttons and sliders, or use gestures to adjust their text without writing additional prompts. In summary: We argue that direct manipulation UIs have the potential to enable writers to produce fewer prompts that more accurately express their intentions, with minimal interruptions, thus protecting their cognitive processes.

(2) AI integration often relies on context switching and fragmented UIs. A second issue arises from the context switching required by recent solutions such as conversational AIs. Writers

seeking AI support must switch away from composing their text when interacting with the AI, often in a separate chat window. This imposes metacognitive demands [8]. Designs integrating AI directly into the writing environment tend to layer AI "on top", adding various UI elements, such as sidebars or pop-ups, which may distract from the central focus: the text itself. This leads to fragmented UIs and thus processes [2]. In contrast, direct manipulation paradigms can integrate AI support into the text UI itself, potentially reducing clutter and minimizing context switches. By enabling writers to interact directly with their writing's UI representation, this approach promises a more cohesive interface with less context switches and lower demands that detract from the writer's own flow of thinking.

To this end, we next review recent work, along with findings from two of our own recent studies.

2 How Direct Manipulation Paradigms Help Design AI Writing Tools as Tools for Thought

For this short position paper, we selected recent examples form an emerging line of research on direct interaction with generative AI.

2.1 Exploring the Impact of Direct Manipulation Interfaces on Peoples' Cognition

A key example is *DirectGPT* by Masson et al. [7], which layers a UI on top of an LLM to map direct manipulation actions to (parametrised) prompts. It provides a visual representation of the objects being edited (text, code, or vector images), helping users maintain a clear mental model and reducing cognitive load. A toolbar to reuse prompting syntax so users can perform tasks without memorizing complex commands, letting them focus on their work rather than on prompts. Outputs can also be manipulated directly; for example, users can drag and drop text elements to reorder them. This makes users' intentions explicit as visual actions which could support them to think more concretely about the changes they make. Undo and redo buttons further encourage people to explore their ideas without the fear of costly mistakes. Indeed, participants were 50% faster and used 50% fewer and 72% shorter prompts compared to baseline ChatGPT. Although cognitive load was not explicitly measured, the reduction in prompt effort likely reduces cognitive demands throughout interaction. This example underscores the potential of how direct manipulation can enhance writing and thinking while preserving cognitive resources.

TaleBrush by Chung et al. [3] uses a line-sketch interface to shape characters' narrative arcs. As users adjust their sketches, the system continuously updates the story, allowing them to better understand and influence AI outputs. This hands-on approach helps users think more concretely about the changes they are making. Although the interaction occurs outside the text itself, line sketches eliminate the need to craft detailed prompts. As shown in the user study, participants were able to use the drawing motions to reliably control their writing without explicit prompts. This interaction further enabled users to generate novel ideas that align well with their intentions. We conclude that by drawing on DM paradigms, *TaleBrush* sketches out how AI writing tools can empower users to express thinking on a higher level of abstraction (here: narrative arcs) without writing prompts.

In recent work, Masson et al. [6] present *Textoshop* a text editing environment that supports direct manipulation via UI elements and instruments inspired from graphics software. For instance, the user can navigate the space of tones in their text using a traditional colour picker. Interaction techniques like this resulted in participants reporting being 14% more successful in accomplishing writing tasks with *Textoshop* compared to a prompting baseline. One participant noted that "[with the baseline] it's quite hard to adjust the prompts and I need [...] several rounds of prompting", underlining that their intentions were implemented more efficient in a less demanding way. The authors also found that it was easy to pick up these concepts quickly, hypothesizing that "participants' knowledge of drawing software likely transferred to *Textoshop*". As a whole, *Textoshop* showcases how direct manipulation can support users – regardless of domain expertise – to leverage AI writing tools without needing to adapt their thinking and workflows to fit the (disruptive) constraints and demands of conversational UIs.

2.2 Content-Driven Local Response

On desktops, users can open CUIs and text editors side-by-side or use an AI sidebar (e.g., *Wordcraft* [9]). Due to limited screen space, this setup doesn't translate well to smartphones. As a result, people writing on mobile devices with AI assistance face frequent context switches between their writing app and a browser or AI app (e.g., ChatGPT). We address this with a new UI concept, Content-Driven Local Response (CDLR) [11]. It enables users to insert responses into the email by directly selecting sentences via touch, which additionally serves to guide AI suggestions. Thus, since the text itself serves as the interface, users can keep the full email in view when interacting with AI, eliminating the need to recall the incoming message. Further, CDLR lets users answer manually or prompt AI with sentences or simple keywords in the same UI location, minimizing workflow interruptions. Our design (cf. Figure 1) supports AI assistance but keeps it optional and encourages users to think about which sentences to respond to before involving AI.

Our study results provide insights into how CDLR may have protected participants' thought processes. We found support for more diverse workflows (usage logs), compared to full reply generation and manual writing. This diversity suggests that participants retained greater agency over their interaction with the system – choosing when and how to engage with AI tools rather than following a fixed path of writing (and thinking). Our interaction logs further suggest that participants used the CDLR principle (replying locally within an email) more often for longer emails. This suggests that CDLR was particularly beneficial for responding to lengthy text, as it reduces the cognitive overhead of remembering the entire content. Overall, CDLR suggests that direct interaction with the text as the primary UI, using small, local, and interactive UI elements, may help protect peoples' thinking and encourages more deliberate use of AI.

2.3 Mobile Touch Interaction with LLMs

In other recent work, we propose to control LLMs via touch gestures performed directly on the text display on smartphones [12]. By using our spread-to-generate gesture (cf. Figure 2), writers can select where and how much text to generate by touching a sentence

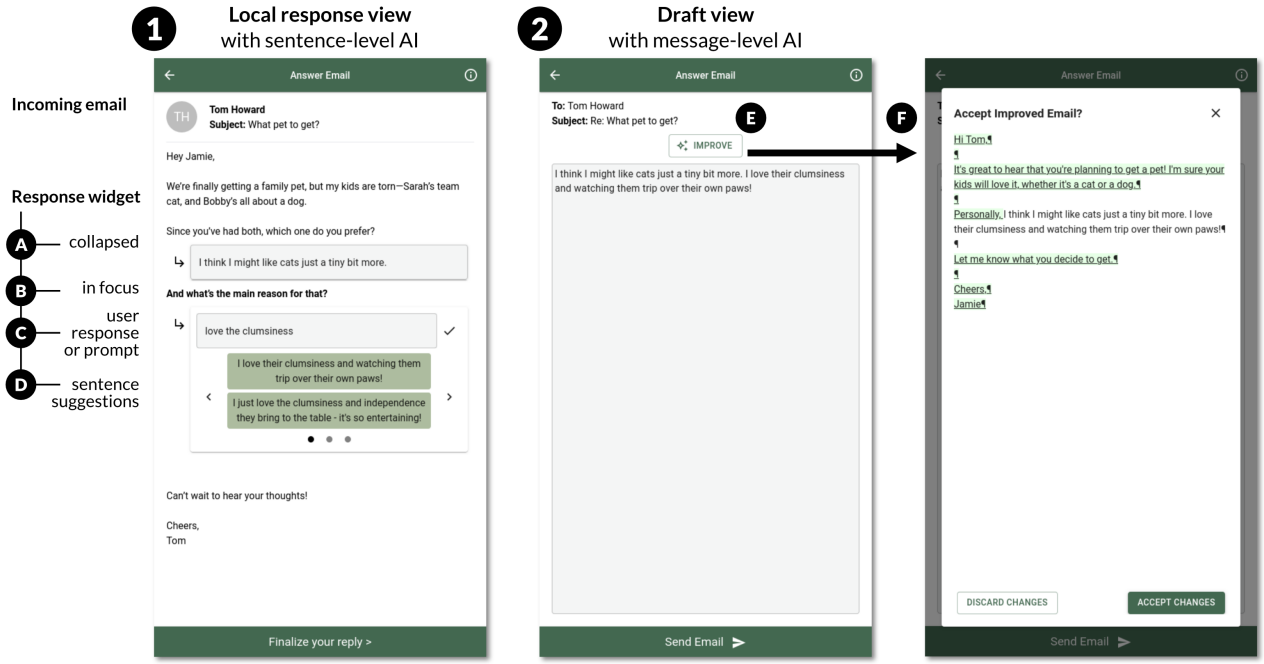


Figure 1: Replying to an email with *Content-Driven Local Response*: *(1)* In the *local response view*, users can insert responses *(A)* directly while reading the email. *(B)* Tapping on a sentence opens a response widget, *(C)* with a text box where users enter a response or a prompt that affects *(D)* the sentence suggestions below. *(2)* After adding local responses, users go to the *draft view*, to turn their responses into a full reply email. They can do so manually and/or with the help of *(E)* an AI improvement pass feature, which generates *(F)* a message-level suggestion, displayed with highlighted changes. These AI features are flexible and optional: Users can add local responses without using suggestions. They can also skip directly to the draft view, optionally enter a prompt there, and use the improvement feature to generate a full reply directly. This supports flexible workflows and allows for direct manipulation.

and adjusting the distance between their fingers. This replaces an explicit prompt and allows users to gradually adjust their intent through continuous interaction, rather than turn-based. New words appear in real time, and users can pause, stop, or regenerate text fluently. To support this, we designed a "Bubbles" visualisation for representing continuous text changes. In this way, people retain direct control over how much AI-generated content enters their document, which respects their own thinking.

This is reflected in our study findings: First, we indeed found that using our DM approach was less mentally demanding compared to a prompting baseline mimicking the ChatGPT UI. Furthermore, participants reported a significantly lower workload (NASA-TLX) and a significantly higher usability (SUS) for the touch interactions compared to the GPT-like interface, which may indicate that DM techniques preserved cognitive resources overall.

Second, when using touch gestures, users were able to complete writing tasks more than twice as fast compared to the prompting baseline. Since participants determined for themselves when their text was complete, this suggests they were able to convey their intentions to the AI more efficiently.

Lastly, participants consistently rated our touch interactions higher in terms of natural feel and control, indicating a more intuitive and streamlined AI writing support system. This reinforces the idea that the DM interface not only enhances efficiency but also preserves users' intentions and thought processes.

3 Discussion

We have presented a short exploration of how direct manipulation paradigms might be used to safeguard and enhance users' cognitive processes in AI-assisted writing.

3.1 Opportunities of Direct Manipulation

Integrating DM paradigms into AI writing tools may reduce the cognitive burden of writing prompts and minimize disruptions to users' thought processes. Here, we summarise our key points of how direct manipulation principles may guide the design of "Tools for Thought":

- Direct manipulation interfaces, as seen in DirectGPT, Textoshop and TaleBrush, may **streamline prompt creation**, *freeing cognitive resources* for higher-level writing and reasoning.
- As a **mediator between user and AI**, direct manipulation and interaction instruments [1] may empower people to *communicate their thoughts and intentions* in a more intuitive, hands-on way, without the demands imposed by writing

Figure 2: Our *spread-to-generate* touch gesture for controlling generative AI on mobile devices. Touches are marked in orange. *(1)* Placing two fingers on the screen sets the cursor (red) to the end of the sentence at the first touch (here: top touch). *(2-3)* Spreading the two fingers fades in blue "word bubbles", which indicate estimates of length and number of words to be generated. In the background, an LLM generates text and streams it to the UI, where it is inserted into empty bubbles as it becomes available. *(4)* Reaching the end of a sentence turns the word bubbles into one green sentence bubble. Further spreading the fingers starts generating another sentence. *(5)* A confirmation widget is shown when lifting the fingers. Tapping the check mark accepts the generated text for *(6)* the final result.

prompts. This may result in UIs that help people shape and refine their own ideas, rather than adapting their thinking to fit AI constraints.

- Direct interaction of **persistent interface objects by both user and AI** mitigates context switching and fragmented workflows and *supports continuity of thoughts*, thus protecting users' cognitive processes.
- Combining **continuous direct interaction with immediate feedback on the AI's responses** helps users evaluate and refine generated content and provides opportunities for control throughout, thus contributing to users' working with their *original thoughts* rather than passively adopting AI outputs.

While the above points establish a clear connection between DM principles and their usability benefits, they also underscore the potential of these principles to support and enhance human cognition. We therefore see these as promising directions for future research into the unique cognitive advantages of DM, that extend beyond mere usability improvements.

3.2 Challenges & Limitations

Although these findings are promising, we recognize potential drawbacks and challenges that could emerge with direct manipulation interfaces. While replacing natural language prompts with UI objects may reduce cognitive demands, it might also make AI usage more obscure to users than, say, an explicit context switch to an AI app. This might amplify undesirable effects of using AI. Similarly, since often writing *is* a part of thinking, writing fewer textual instructions could lead to users thinking *less* about their intentions for involving AI. Lastly, DM blends user input and model output at the same UI objects, in contrast to, e.g., separated speech bubbles in CUIs. If this mixing takes place within the text, AI-generated content might also bleed into human thinking, as related work on opinionated language models suggests [5].

Looking ahead, our findings here are limited to reviewing recent studies of direct manipulation in AI-assisted writing, where cognition was not the primary focus. Therefore, we only reflect on the design choices in this recent work and their outcomes to highlight *potential* cognitive impacts, to be examined in detail in the future.

4 Conclusion

In this position paper, we advocate for further exploring direct manipulation as a means to protect and augment human thinking – concretely by reducing disruptions caused by writing out prompts, potentially mitigating cognitive overhead, and helping writers maintain their flow of thoughts. However, more research is needed to investigate how specific direct manipulation techniques affect different aspects of cognition, cognitive load, ideation, and user agency, across various writing tasks. Moving forward, research on these topics should also address both the ethical considerations and broader societal implications of widespread adoption of generative AI tools for writing and thinking.

References

[1] Michel Beaudouin-Lafon. 2000. Instrumental interaction: an interaction model for designing post-WIMP user interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (The Hague, The Netherlands) *(CHI '00)*. Association for Computing Machinery, New York, NY, USA, 446–453. https://doi.org/10.1145/332040.332473

[2] Daniel Buschek. 2024. Collage is the New Writing: Exploring the Fragmentation of Text and User Interfaces in AI Tools. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference* (Copenhagen, Denmark) *(DIS '24)*. Association for Computing Machinery, New York, NY, USA, 2719–2737. https://doi.org/10.1145/3643834.3660681

[3] John Joon Young Chung, Wooseok Kim, Kang Min Yoo, Hwaran Lee, Eytan Adar, and Minsuk Chang. 2022. TaleBrush: Sketching Stories with Generative Pretrained Language Models. 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 209, 19 pages. https://doi.org/10.1145/3491102.3501819

[4] Hai Dang, Sven Goller, Florian Lehmann, and Daniel Buschek. 2023. Choice Over Control: How Users Write with Large Language Models using Diegetic and Non-Diegetic Prompting. 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 408, 17 pages. https://doi.org/10.1145/3544548.3580969

[5] 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. https://doi.org/10.1145/3544548.3581196

[6] Damien Masson, Young-Ho Kim, and Fanny Chevalier. 2024. Textoshop: Interactions Inspired by Drawing Software to Facilitate Text Editing. arXiv:2409.17088 [cs.HC] https://arxiv.org/abs/2409.17088

[7] Damien Masson, Sylvain Malacria, Géry Casiez, and Daniel Vogel. 2024. DirectGPT: A Direct Manipulation Interface to Interact with Large Language Models. In *Proceedings of the 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. https://doi.org/10.1145/3613904.3642462

[8] Lev Tankelevitch, Viktor Kewenig, Auste Simkute, Ava Elizabeth Scott, Advait Sarkar, Abigail Sellen, and Sean Rintel. 2024. The Metacognitive Demands and Opportunities of Generative AI. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 680, 24 pages. https://doi.org/10.1145/3613904.3642902

[9] Ann Yuan, Andy Coenen, Emily Reif, and Daphne Ippolito. 2022. Wordcraft: Story Writing With Large Language Models. In *Proceedings of the 27th International Conference on Intelligent User Interfaces* (Helsinki, Finland) *(IUI '22)*. Association for Computing Machinery, New York, NY, USA, 841–852. https://doi.org/10.1145/3490099.3511105

[10] J.D. Zamfirescu-Pereira, Richmond Y. Wong, Bjoern Hartmann, and Qian Yang. 2023. Why Johnny Can't Prompt: How Non-AI Experts Try (and Fail) to Design LLM Prompts. 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 437, 21 pages. https://doi.org/10.1145/3544548.3581388

[11] Tim Zindulka, Sven Goller, Florian Lehmann, and Daniel Buschek. 2025. Content-Driven Local Response: Supporting Sentence-Level and Message-Level Mobile Email Replies With and Without AI. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '25)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3706598.3713890

[12] Tim Zindulka, Sekowski Jannek, Florian Lehmann, and Daniel Buschek. 2025. Exploring Mobile Touch Interaction with Large Language Models. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '25)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3706598.3713554