Beyond chat: collaborative editors enhance human involvement and agency when co-writing with large language models

Rodolfo Ocampo
The University of New South Wales
Creative Technologies
Research Lab (CTRL)
Sydney, NSW, Australia
r.ocampo_blanco@unsw.edu.au

Oliver Bown
The University of New South Wales
Creative Technologies
Research Lab (CTRL)
Sydney, NSW, Australia
o.bown@unsw.edu.au

Kazjon Grace
The University of Sydney
Design Lab
Sydney, NSW, Australia
kazjon.grace@sydney.edu.au

Abstract

As artificial intelligence (AI) tools become increasingly common for assisting with writing tasks, chat-based interfaces have emerged as the dominant mode of interaction. We hypothesise that relying solely on chat interfaces reduces users' involvement in the writing process, relegating them to an instructing role rather than an active co-author. To investigate this, we developed two prototype systems with hybrid interfaces integrating a *chat window* next to a *collaborative writing editor* where both human and AI can contribute, and we compared them to *chat-only* systems. Across two user studies, participants who had access to the collaborative hybrid interface reported higher levels of personal contribution, perceived a more balanced partnership with the AI, and adopted more engaged roles than participants using the chat-only interface. These findings suggest that offering a collaborative editor alongside chat can mitigate the risk of overreliance on AI and help preserve creative skills. As generative AI tools become more pervasive, designing interfaces that maintain user involvement is crucial to supporting meaningful human–AI co-creation.

Introduction

In recent years, advances in large language models (LLM) have made writing with them commonplace (Sun et al. 2024; Palani and Ramos 2024; Chakrabarty et al. 2024). Despite clear benefits such as overcoming writer's block and supporting non-native speakers, concerns are emerging about overreliance and loss of skill (Gerlich 2025; Lee 2025).

Users may allow the model to produce entire texts without engaging in thorough proofreading or revision, leading to generic-sounding outputs, factual errors, and a gradual erosion of writing and critical thinking skills (Gerlich 2025; Abbas, Jam, and Khan 2024; Essel et al. 2024; Heersmink 2024). Furthermore, excessive automation can diminish the intrinsic satisfaction emerging from creative work. As Brian Eno points out when discussing the limitations of generative AI; much of the value in creative activities lies in the process rather than in the final outcome (Eno 2024).

To mitigate these risks, it is crucial to design co-creative AI systems that actively encourage user involvement. This is largely an interaction and interface design problem. Currently, chat-based interfaces have emerged as the dominant paradigm for interacting with LLMs. This approach, however, may position users as mere requesters who delegate writing tasks to the AI rather than as writing co-creatos. In other words, chat interfaces provide primarily a space to interact *about* the writing, but not *through* the writing, both important in co-creation (Bown et al. 2020a; Rezwana and Maher 2022a; McGuire, De Cremer, and Van de Cruys 2024).

To address this issue, we hypothesise that implementing **hybrid interfaces with both a chat window to discuss the writing and a collaborative editor to contribute to it**, can increase user involvement.

We investigated this hypothesis through two studies, both employing hybrid interface prototypes of increasing fidelity. In the first study, 23 participants were divided into two groups: one used the hybrid collaborative interface and the other the chat-only interface. Compared to those using the chat-only interface, participants who used the collaborative editor were less likely to report that the AI did most of the work and more likely to report the final piece contained their own words. Moreover, they described their role as more involved in the writing, while participants in the chat-only interface were more likely to describe their role as providing ideas, instructions and requests to the AI.

Building on these findings, we conducted a second study with 25 participants using an improved hybrid prototype. This follow-up was not comparative: instead it focused on collecting qualitative in-depth feedback about user experience, though it included questions prompting users to compare it with their previous experience with popular tools like ChatGPT. Results indicated that the shared text editor led to reduced friction, an enjoyable experience, and fostered greater user involvement with the writing process. However, participants identified several key limitations: they struggled to clearly identify which edits were made by the AI, requested better version control and revision history features, and suggested adding GUI elements that could

trigger premade actions to streamline their workflow.

Related Work and Background

Large Language Models (LLMs) have evolved considerably in their interaction paradigms. Initially developed as autocomplete systems (Vaswani et al. 2017; Brown et al. 2020), they transitioned to instruction-based interaction with InstructGPT (Ouyang et al. 2022), before adopting the now ubiquitous chat interfaces. ChatGPT's rapid adoption, becoming the fastest application to date to reach 100 million users (Heaven 2023), stemmed not from novel technology but largely from the transition to a dialogic chat interface. As OpenAI's President Greg Brockman explained: "we actually had the model [GPT3.5] behind it created almost a year prior so it wasn't new technology, the thing that we really did differently is that we did a little bit of extra work to make it more aligned so you could talk to it and it would do what you wanted". This interface allowed them to bridge "what people thought was possible and what actually had been possible for quite some time" (SXSW 2023). Jan Leike, another ChatGPT contributor, when explaining reasons for its rapid adoption, claimed: "It's not a fundamentally more capable model than what we had previously. The same basic models had been available on the API for almost a year before ChatGPT came out [...] we made it more aligned with what humans want to do with it. It talks to you in dialogue, it's easily accessible in a chat interface" (Heaven 2023).The evolution toward chat-based interaction has been followed by numerous systems including Gemini, Llama, and Claude (Touvron et al. 2023; Gemini Team 2024).

While chat interfaces enhanced accessibility, they potentially limited creative collaboration by separating users from direct engagement with the creative artifact. Earlier autocomplete paradigms allowed interaction *through* the writing itself, whereas chat interfaces shifted to interaction *about* the writing. This distinction of types of interactions has been studied in the literature by some authors (Calderwood et al. 2020) and having distinct spaces for discussion and co-creation has been identified as important for successful co-creation (Bown et al. 2020b; Rezwana and Maher 2022a; Coenen et al. 2021). We are particularly interested in understanding how this hybrid interface affects human direct involvement and user experience, particularly in the context of highly capable language models.

Human-AI Co-Creativity and Mixed Initiative Systems

Mixed-initiative interaction enables both human and computers to actively contribute to tasks fostering user agency through iterative exchange (Allen, Guinn, and Horvtz 1999; Horvitz 1999). When applied to creative domains, this creates mixed-initiative co-creativity, where humans and AI contribute as equal collaborators rather than in client-producer relationships (Muller, Weisz, and Geyer 2020; Yannakakis, Liapis, and Alexopoulos 2014; Davis 2013).

Mixed-initiative co-creative interfaces differ from request-based systems by establishing a tightly interactive loop where both parties suggest, produce, evaluate, and modify outputs within a shared creative space. Examples include collaborative environments for drawing (Karimi et al. 2020) and writing (Kantosalo et al. 2016). Research shows that effective co-creativity requires both a shared creation space and communication channels (Rezwana and Maher 2022b), with communication establishing the common ground necessary for collaboration (Dafoe et al. 2021).

Bown et al.'s framework of "dialogic creative interaction" distinguishes between interacting *through* an artifact (directly modifying it) and *about* an artifact (exchanging ideas regarding creative intent) (Bown et al. 2020b). Similarly, Rezwana and Maher differentiate between interactions *between* collaborators and *with* the creative product (Rezwana and Maher 2022a). Conventional chat interfaces conflate these interaction modes, constraining collaborative possibilities and limiting the capacity for stateful co-creation.

Shneiderman's direct manipulation principles, emphasising controllability, predictability, and continuous representation, further highlight chat systems' limitations for co-creativity (Shneiderman 1997).

Beyond Chat Interfaces for Co-Creative Generative AI Systems

Recent systems have begun distinguishing between communication spaces and co-creation spaces. Anthropic's Claude introduced "Artifacts" (Anthropic 2024), though as of September 2024, only the AI can modify these elements. ChatGPT's Canvas feature and specialized writing tools like Sudowrite, Jarvis, and Wordcraft (Yuan et al. 2022) provide shared spaces where both human and AI can edit text while communicating through separate channels. Similar developments appear in coding environments, with systems like Cursor (Cursor 2024) and GitHub Copilot enabling both communication and collaborative editing, though with varying degrees of integration.

While multiple studies have examined human co-writing with LLMs from different perspectives (Wan et al. 2023; Yuan et al. 2022; Hwang et al. 2024; Chakrabarty et al. 2024; Grigis and De Angeli 2024; Masson et al. 2024; McGuire, De Cremer, and Van de Cruys 2024; Lee et al. 2024), our contribution is **specifically exploring the effect that collaborative spaces in addition to chat interfaces** may have on human-AI co-creativity, focusing on user agency, involvement, immersion, enjoyment, creative self-efficacy and perceived roles. To our knowledge, no such evaluations exist in the literature.

Moreover, the studies and prototypes presented here predate the recent introduction of collaborative spaces such as Artifacts in Claude and Canvas in ChatGPT. As such, the insights presented in this paper are timely to

understand the co-creative potential and limitations of these interfaces.

Prototypes

To investigate the efficacy of hybrid chat-and-editor interfaces, we developed two sequential prototypes that implemented this functionality. This iterative development process yielded first Vorges, followed by a refined version named Common AI. Both systems were deployed as web applications built using Node.js, with the OpenAI API providing the underlying large language model capabilities.

Version One – Vorges

Vorges, developed in 2023, represented our initial exploration into human-AI collaborative writing environments. The system utilised the GPT-3.5 model via the OpenAI API and implemented an agentic approach wherein the LLM was designed both to communicate with users and to directly manipulate a shared text editor in the front end.

We established a rudimentary communication protocol between the front-end application and the language model through an API, which enabled the AI to perform two distinct functions:

- Communicate with the user through the chat interface
- Modify content within the collaborative text editor

To facilitate text editing, the LLM was instructed to enclose editor text modifications within curly brackets, with content outside these delimiters appearing in the chat window. This parsing logic was managed by the application back-end.

A significant limitation of this implementation was its requirement for the AI to return the entire edited text within curly brackets, even for minor modifications. This wholesale replacement approach resulted in increased computational costs, saturation of context windows, and from a co-creative perspective, diminished ability to perform fine-grain edits efficiently.

Version Two – Common AI

The second iteration, Common AI (Figure 1), developed in late 2024, represented a substantial refinement of our approach. This prototype leveraged the more sophisticated GPT-4o model and implemented a structured JSON protocol for granular text manipulation. This advancement enabled the AI to execute precise edits without necessitating resubmission of the entire document.

The enhanced protocol expanded the AI's editing capabilities to include:

- Appending text at the beginning or end of the document
- Editing specific textual segments, from individual words to comprehensive revisions
- Deleting targeted text from the document
- Communicating with the user through the chat interface

This protocol enhancement improved system usability while optimising the management of context windows, ultimately enabling more fine-grained collaboration between the user and the LLM.

User Studies: Designing a Shared Creative Workspace and Examining its Effect on Human-AI Co-Creative Writing

We conducted two exploratory studies. The first one was conducted in June 2023 using the first version of the prototype (Vorges) while the second one, using an improved version (Common AI), was conducted in August 2024. These investigations sought to understand how implementing a shared collaborative space alongside a chat window might enhance user involvement and creative agency in human-AI collaboration.

Study 1

Context The initial exploratory study employed the Vorges prototype and was conducted in situ at Sydney's Powerhouse Museum during an AI-themed public event. Participants were recruited through a combination of opportunistic sampling as visitors explored the museum and direct attendance following promotional materials in the event program.

Experimental Design We established six computer stations across three tables, implementing a between-subjects comparative design. Three stations featured the Vorges interface with both collaborative editor and chat window, while three control stations offered a chat-only version of the same system. This configuration allowed direct comparison between the two interaction modalities within the same experimental context.

Participants The study included 23 participants (age range: 18–53 years) who were randomly assigned to available stations. This resulted in 12 participants using the chat-only interface and 11 using the chat-and-editor interface.

Task Participants were asked to write a micro-science fiction story set in a speculative post-2050 future. The story had to comprise at least three paragraphs, with no imposed time constraints.

Measures Our investigation focused on comparing experiences between the two interface conditions across five key dimensions: User involvement and agency; Sense of collaboration; Enjoyment and satisfaction; Creative self-efficacy; Emergent user-AI roles

After task completion, participants completed a mixed-methods survey containing open-ended qualitative questions and Likert-scale quantitative items.

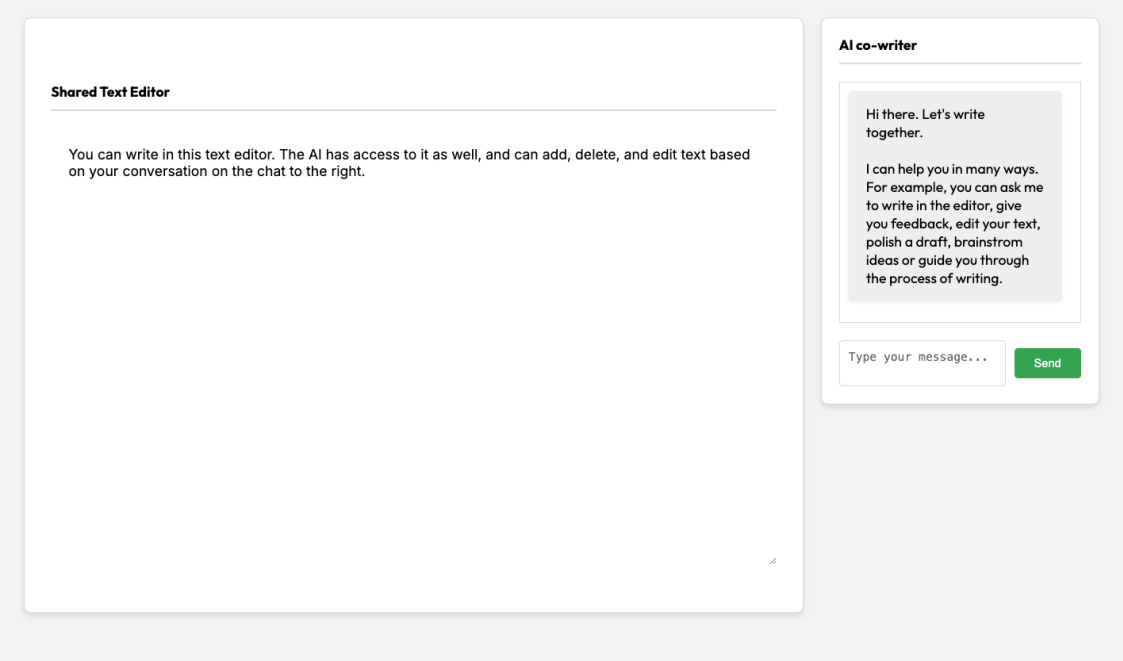

Figure 1: Screenshot of Common AI, our implementation of a hybrid interface, integrating a collaborative editor and a chat window. The previous version Vorges, had a similar interface, but a less robust back-end functionality, as well as a dark theme. Common AI's improved capabilities allowed the AI to perform fine-grained document modifications and utilised a more capable underlying language model. Common AI served as the experimental platform for the second study presented in this paper.

Methods and Analysis Our analysis of user responses was guided by Creswell's (Creswell 2012) qualitative research approach, particularly narrative research, used to understand and contextualise user's personal experiences and meaning-making. Moreover, qualitative responses underwent thematic analysis to identify emergent patterns (Braun and Clarke 2012). Likert-scale responses were converted to numeric values (1=Strongly Disagree to 5=Strongly Agree) for statistical comparison. We employed t-tests to assess statistical significance between mean responses for each question across the two conditions.

Study 2

Context The second study utilized the refined Common AI prototype and was conducted online, allowing participants to access the collaborative environment remotely through a web interface. This study aimed to gather deeper insights into user experiences with the shared text editor.

Participants We recruited participants (n=25) through social media channels and snowball sampling techniques. Participants received incentives in the form of $15 AUD vouchers (Westfield or Uber Eats). Eligibility required prior experience using chat-based LLM tools (e.g., ChatGPT, Claude) for writing tasks.

Study Design This investigation employed a single-condition exploratory design focused on the chat-and-editor interface. Given the sample size limitations, which precluded robust quantitative statistical analysis, we prioritized gathering rich qualitative feedback to illuminate the key challenges and possibilities within the collaborative interface.

Task Participants selected one of five creative writing prompts, which they completed and submitted with no time constraint.

Methods and Analysis After task submission, participants completed a mixed-methods survey emphasising open-ended responses. The instrument sought to understand participants' experiences with the collaborative editor, its perceived limitations, and comparative reflections against their previous experiences with other writing tools. Similar to study 1, we utilised a narrative research approach coupled with thematic analysis (Braun and Clarke 2012; Creswell 2012). Additional Likert-scale items measured perceived levels of user and AI involvement in the creative process.

Study 1 Results

Differences in work distribution and involvement across chat-only and hybrid collaborative editor interfaces

Our initial exploration revealed differences across interface modalities, as shown in Figure 2. While the Likert-scale survey did not achieve statistical significance due to a low sample size, differences between interfaces and questions indicated that the most marked differences where in attribution of who did most of the work, and levels of immersion and direct involvement with the writing. Participants in the collaborative editor + chat interface were more likely to report the final story contained words they wrote and reported higher levels of immersion in the writing task. Conversely, participants in the chat-only interface predominantly attributed the writing work to the AI system (Vorges), more strongly agreeing that the "AI did most of the work".

Qualitative Results: Understanding Perceived and Assumed Roles in Each Group The qualitative data revealed nuanced differences in how participants conceptualised the role they played in the creative process compared to the role that AI played. While both groups generally reported relying on Vorges to contribute substantially to the writing, the nature and degree of this reliance varied between conditions.

Comparative Role Dynamics Between Interfaces In the chat-only interface, participants predominantly positioned themselves as idea providers while ceding the writing process to the AI. This dynamic is exemplified by participant 9, who simply stated: "Vorges generated all of the text," and participant 6, who when asked about Vorges' role, responded with a single word: "Everything." A more elaborated response from participant 11 illustrates this delegation of creative labour: "I mainly just gave the ideas and instructions, such as: write a paragraph describing the character as x, y, z. Vorges nicely narrated my ideas. It was like giving the recipe of a cake to someone and letting it do it."

In contrast, participants using the collaborative editor + chat interface were more likely to report starting with, or at some point doing some of the writing, and then using the AI to improve or expand it. Participant 3 explained: "I wrote a simple story, easy language, not very detailed, and then I asked Vorges to make my story more advanced, bigger words, more description, more intense." Similarly, participant 4 noted that the AI "took the story in fun and interesting directions, and added to the experience rather than entirely controlling the experience".

Contrasting Perceptions of AI Contribution The differential framing of the AI's contribution was another revealing pattern. Chat-only interface users predominantly characterised Vorges as helping to *write their ideas*, whereas collaborative editor + chat users described the AI as helping them *ideate* or *improve their own writing*. This subtle distinction reflects fundamentally different conceptualisations of creative ownership and process.

User Satisfaction and Sense of Accomplishment This divergence in experience extended to satisfaction with the process itself. One participant from the chat-only condition (participant 11) expressed a notable ambivalence: "it also made me feel like I was cheating somehow. It does not feel like my work, even though I gave all the ideas. Also, I believe there is satisfaction in putting a lot of effort/dedication/patience into something. Vorges made everything so simple, fast, and easy that it felt artificial and no real satisfaction came as a result." This sentiment contrasts sharply with reports from the collaborative editor + chat group, where participant 7 described the process as "fun and fulfilling," and participant 2 expressed enthusiasm about future engagement, stating they "Would love to do it again."

Challenges and Limitations of the Collaborative Interface However, users also reported challenges with the collaborative editor + chat interface. Participant 3, who used Vorges to enhance their writing with more sophisticated language, noted: "I still wrote my own story, but the style of the writing no longer feels like mine." This highlights tensions between enhancement and authenticity that merit further exploration. Additionally, participant 9 found the system "not intuitive".

These initial findings prompted us to develop an enhanced prototype for Study 2 addressing technical challenges of the previous prototype and seeking to achieve more in deeper insights about the unique challenges and potential of implementing a collaborative editor.

Results Study 2: Refined Interface and Deeper Qualitative Insights

Overview Study 2 employed an improved prototype leveraging a more sophisticated model capable of executing fine-grained text edits. Rather than replacing entire text blocks to accommodate even minor changes, the enhanced system (Common AI) could perform targeted edits: adding text at specific locations, deleting selected words or sections, and making punctual modifications.

This study was aimed at gathering richer qualitative insights through predominantly open-ended questions. Unlike the comparative approach of Study 1, this follow-up investigation focused on an in-depth qualitative exploration of user experiences with the improved collaborative editor + chat interface.

Additionally, we sought to further investigate a key finding from the previous study: the tendency of chat-only interface users to attribute most of the creative work to the AI. To this end, we incorporated Likert scale questions assessing agreement with the statements "The AI did most of the work" and "I did most of the work."

Importantly, this second study was conducted in a significantly different wider technological context than

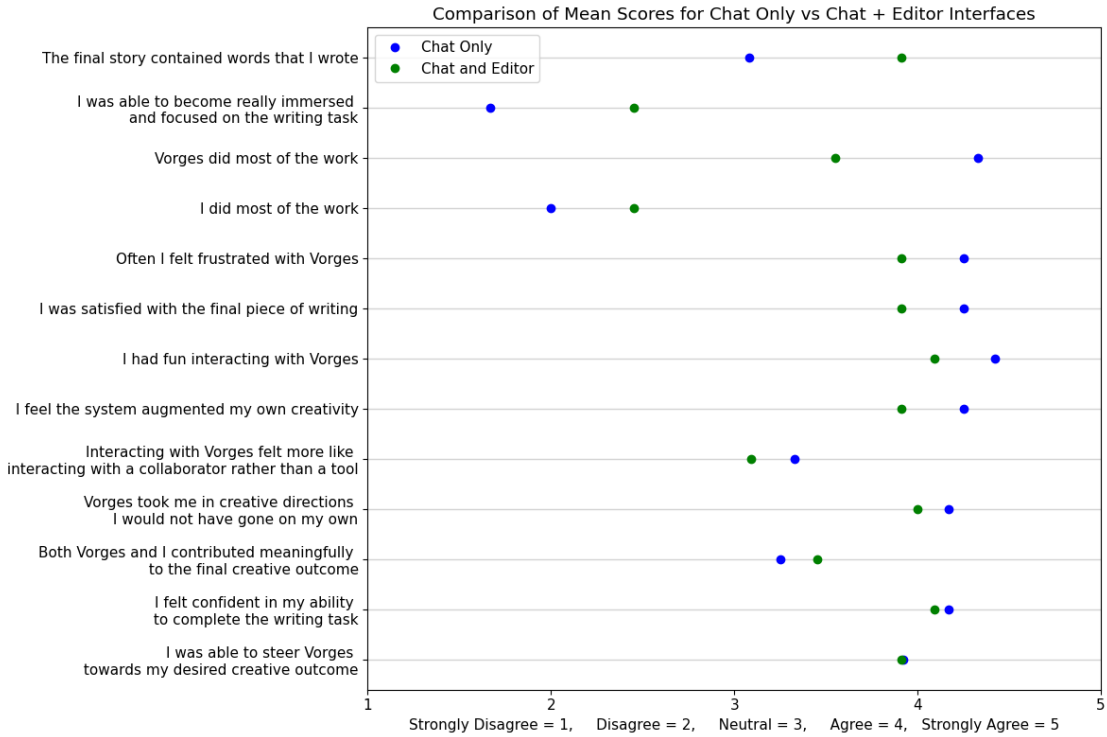

Figure 2: Comparison of mean scores for the Likert Scale questions. Out of the 13 questions, 9 yielded small differences while 4 yielded larger observable differences, showing, generally, that users of the collaborative editor were more involved and immersed in the writing, while users of the chat-only interface thought the AI did most of the work

Study 1, as chat interfaces and AI systems like ChatGPT, Claude, and Gemini had become ubiquitous in the interim. This shifting landscape allowed us to explore user frustrations with existing AI writing tools and position our findings within a more mature ecosystem of LLM-based systems.

High Enjoyment and Involvement, with an Expressed Need to Highlight AI-Generated Changes Participants generally reported positive experiences using the tool, with 20 out of 25 participants describing highly enjoyable interactions. For example, Participant 4 shared: "I really-really enjoyed writing this. I even had a deep moment of reflection, my writing was nostalgic and sad, but I was able to use AI to steer it in the right direction, it gave me confidence that I was also writing with correct grammar and spelling, English is not my first language and while I am proficient, I can still use proofreading to ensure good quality, this tool helped me with it." This participant primarily engaged with the editor component and sought AI assistance to enhance their writing. When asked about limitations, they noted: "I am unsure as what parts of the text are being edited, in the end I am not quite sure which parts were mine and which ones were edited by AI."

A similar concern was expressed by Participant 7, who found the system "very helpful, helps me bring my ideas to life in a more vivid and captivating way," but identified a limitation: "Once the tool has made revisions to the original text, maybe it can highlight the key changes that have been made, that way I can easily see where the tool has made updates, otherwise I need to slowly read through and identify the changes myself."

Participant 11 reinforced this theme while describing their editing-focused workflow: "When you ask the AI to check your grammar (as I did), it would be good if it told me what suggestions it had made, so I can double check its work easier."

These reflections consistently highlight users' desire to maintain clarity about the boundary between their contributions and AI-generated content—suggesting a prioritisation of collaborative transparency that preserves their creative agency rather than amalgamating human and AI work indistinguishably.

Emphasising the importance of maintained agency, Participant 9 noted: "I liked how my original ideas were still retained and AI was used as a tool to complement my intentions. It forced me to put in some effort and do the majority of the work, I enjoyed how I could not be fully reliant on AI to provide me with all the text."

Interestingly, all participants who expressed these positive experiences and desire for clear attribution also reported that they had done most of the work when responding to the Likert scale questions, disagreeing with the statement that "The AI did most of the work." and

agreeing with the statement "I did most of the work"

Users Who Outsourced Most of the Work Found Limitations with Writing Style A contrasting pattern emerged among users who reported that the AI had performed most of the work. For instance, Participant 16, who strongly agreed with the statement "The AI did most of the work" and strongly disagreed with "I did most of the work," identified the main limitation as "The writing was not very good" and claimed that having a collaborative text editor "made no difference." Similarly, Participant 21, who also strongly agreed that the AI did most of the work, stated that the collaborative editor "Doesn't make a difference" and expressed that they "don't like using AI to generate actual pieces of writing."

This finding suggests that for users who are less involved at the writing level found the quality of the AI's writing more of a problem than those users who engage, and thus shape the style of the output.

Users Felt the Collaborative Editor Contributed to Higher Creative Agency A recurring theme associated with positive experiences was an enhanced sense of agency and genuine collaboration. Participant 14 articulated this clearly: "Makes sense to me. So much of popular discourse around human and AI interaction centres around collaboration and partnership in creativity. My view is that AI needs human feel to serve the purposes we need it for, i.e. to be human. A shared text editor is a useful format to encourage collaboration and achieve human guidance."

This sentiment was echoed by Participant 12, who appreciated that the system "still let me have autonomy," while Participant 6 valued the "flexibility and control over the final result."

The ethical dimensions of human-AI creative collaboration were highlighted by Participant 22, who noted: "It adds an element of working together, which I think is the moral problem with current AI tools - they often seem like they're doing all the work."

Participant 25 expressed particular satisfaction with the AI's complementary rather than dominant role: "I absolutely love it! I enjoy creative writing, and AI expands my vocabulary and corrects my grammar. I love that it provides feedback before making changes."

Greater Autonomy and Collaboration than Popular Chat-Based Tools When comparing the collaborative editor to conventional chat-based AI tools, participants consistently noted advantages in agency and workflow integration. Participant 18 remarked: "It was much better than ChatGPT I enjoyed how it gave me a lot more agency." Participant 5 highlighted the practical workflow benefits: "it can just be a bit clunky having a separate document to then copy, paste, and edit in. This made it super seamless being in the one program."

Participant 10 expressed a clear preference for the dual-interface approach: "I think it's better than just having one textbox [the chat box]; I like it quite a lot, actually." This sentiment was reinforced by Participant 1, who described "The shared text editor is a great feature and a significant advantage" compared to chat-only tools, and by Participant 17, who independently stated: "I think this is a better solution than typical AI chatbots."

When asked about frustrations with conventional chat-based tools like ChatGPT and Claude, participants frequently mentioned challenges with maintaining control and iterative refinement. Participant 17 noted: "ChatGPT will always rewrite the entire passage to change just one paragraph, and it's harder to work on one text because I often need to scroll back up or continually copy and paste it." Similarly, Participant 15 expressed frustration that "it's gone in the wrong direction and requires a lot of input from me to get it back on track."

Participants Valued the Chat Window Despite Criticism of Its Size An interesting pattern emerged regarding the chat component of the interface. While only one participant requested a larger editor space, multiple users identified the chat window's small size as a limitation -suggesting a continued desire for communication with the AI alongside direct editing.

Participant 17 noted: "The AI chat window felt a little too small, so I had to scroll to see what I had asked it, even within the one prompt." This was echoed by Participant 14 who simply stated: "The chat box was a bit small," and Participant 5 who elaborated: "I found the box on the right rather small. It was also difficult to scroll up to see what you have written in the text box before submitting it."

This emphasis on the chat window's importance, despite its seemingly secondary role to the editor, suggests that a good chat experience and interface remains crucially important for users.

Interface Clarity and Action Affordances as Areas for Improvement A recurring theme in participant feedback concerned the need for clearer signalling of interface components and their functions. Participant 15 explained: "I didn't understand the format of the site at first as the instructions within the big collaborative text area on the left were confusing (and weren't in a different font or distinguishably different to the area where the collaborative story was updating)." They further noted mismatched expectations: "I thought that it would take what I was writing and make more of a change to it by combining with more flourished details about the places I was writing about (e.g. it would come up with details about the landscapes I was in) but it kind of just copied my words and made slight grammatical changes."

Participant 9 simply described the interface as "not that intuitive," while Participant 22 suggested visual design improvements: "make the fields different colours to highlight the difference in the two, have the AI prompt part set out as more of a chat, like messenger, so that the text editor is seen as the output field."

Several participants recommended implementing visible action affordances rather than requiring written requests for common functions. Participant 22 suggested: "Some suggestive prompts would also be useful like ChatGPT, because it is hard to know what the AI is even capable of and that limits the creative output." Despite this critique, they maintained: "I quite like it, it's very different to other AI tools. Definitely takes some getting used to but I absolutely see future applications and usefulness."

In a similar vein, Participant 17 proposed: "Adding some ready made suggested prompts based on my writing, or just in general," while Participant 9 independently suggested "a drop down of up to 3 suggestions based off the prompts." Participant 4 envisioned "a menu to suggest common writing tasks or styles" and noted that "If used for educational settings, scaffolding could help create a more hands-on experience."

Highlighting a desire for more automated assistance, Participant 24 noted: "I had to actively input each command eg grammar and spellcheck. I feel this should be done automatically."

Finally, regarding version control and the visibility of AI contributions, Participant 2 suggested: "Maybe having a revert AI contribution button and/or contribution history so you could go back to past historical copies." While this functionality was actually available through an undo button, the comment suggests that its presence was not sufficiently evident. As Participant 19 noted: "I enjoyed the fact that you can go back to previous revisions of our work, like flipping a page back." However, as previously discussed by other participants, enhanced visibility of specific changes between versions remains an area for improvement.

Discussion and Conclusion

This research investigated how interface design affects human involvement in AI co-creative writing. We posited that conventional chat interfaces may limit human agency, potentially relegating users to mere instruction-givers rather than active co-creators. To address this limitation, we developed and evaluated a prototype combining collaborative editing with chat functionality, enabling users to interact both *through* and *about* the writing process.

Through two iterative studies, we explored this interface's impact on agency, involvement, satisfaction, and creative self-efficacy. Four key findings emerged:

Differential Attribution of Creative Labor Our first study revealed a clear distinction in how users perceived creative contribution across interfaces. In the chat-only condition, participants consistently reported that "the AI did all the work" and noted minimal inclusion of their own writing in the final product. Users primarily described their role as providing ideas while the AI performed the actual writing. Conversely, collaborative editor users reported greater personal contribution to the writing process.

Enhanced Creative Agency The second study, using an improved version of the collaborative editor, demonstrated that users valued maintaining active involvement in the writing process. Participants reported using the AI as a refinement tool to polish and expand their writing rather than surrendering creative control entirely. This approach preserved their sense of agency while still benefiting from AI assistance.

Advantages Over Chat-Only Interfaces Participants explicitly contrasted their experience with conventional chat interfaces, highlighting frustrations with traditional approaches. They noted difficulties in maintaining control over text when AI rewrites entire passages for minor changes and the inefficiency of copying text between systems. The integrated approach was described as "super seamless," offering "more agency" and "flexibility and control over the final result."

Interface Improvement Recommendations Users suggested several enhancements to the collaborative editor experience. These included: highlighting AI versus human contributions, implementing more robust version control, providing clearer indicators of AI capabilities, and offering GUI components for triggering predefined actions. These suggestions align with established HCI principles, including Norman's visibility of affordances (Norman 2013), Shneiderman's direct manipulation concept (Shneiderman 1997), and Amershi's guidelines for human-AI interaction (Amershi et al. 2019).

Our findings, while derived from relatively small participant samples (23 and 25 respectively), provide valuable qualitative insights into the potential of collaborative editing interfaces to enhance human-AI co-creativity. As major platforms increasingly implement variations of collaborative editors alongside chat functionality, understanding how these interfaces affect user involvement becomes crucial for maximising human creative agency in AI collaboration.

Future research should focus on implementing improved prototypes that provide greater clarity about co-creative capabilities, robust history and versioning features, and clear visibility of user versus AI contributions in the collaborative writing process (Haase, Djurica, and Mendling 2023).

References

Abbas, M.; Jam, F. A.; and Khan, T. I. 2024. Is it harmful or helpful? examining the causes and consequences of generative AI usage among university students. *International Journal of Educational Technology in Higher Education* 21(1):1–22.

Allen, J. E.; Guinn, C. I.; and Horvtz, E. 1999. Mixed-initiative interaction. *IEEE Intelligent Systems and their Applications* 14(5):14–23.

Amershi, S.; Weld, D.; Vorvoreanu, M.; Fourney, A.; Nushi, B.; Collisson, P.; Suh, J.; Iqbal, S.; Bennett, P. N.; Inkpen, K.; Teevan, J.; Kikin-Gil, R.; and Horvitz, E. 2019. Guidelines for human-AI interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM.

Anthropic. 2024. What are artifacts and how do I use them? https://support.anthropic.com/en/articles/9487310-what-are-artifacts-and-how-do-i-use-them. Accessed: 2024-9-11.

Bown, O.; Grace, K.; Bray, L.; and Ventura, D. 2020a. A speculative exploration of the role of dialogue in human-ComputerCo-creation. *ICCC*.

Bown, O.; Grace, K.; Bray, L.; and Ventura, D. 2020b. A speculative exploration of the role of dialogue in human-ComputerCo-creation. In *ICCC*, 25–32. computationalcreativity.net.

Braun, V., and Clarke, V. 2012. Thematic analysis.

Brown, T. B.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A.; Agarwal, S.; Herbert-Voss, A.; Krueger, G.; Henighan, T.; Child, R.; Ramesh, A.; Ziegler, D. M.; Wu, J.; Winter, C.; Hesse, C.; Chen, M.; Sigler, E.; Litwin, M.; Gray, S.; Chess, B.; Clark, J.; Berner, C.; McCandlish, S.; Radford, A.; Sutskever, I.; and Amodei, D. 2020. Language Models are Few-Shot Learners. *arXiv:2005. 14165 [cs]*.

Calderwood, A.; Qiu, V.; Gero, K.; and Chilton, L. B. 2020. How novelists use generative language models: An exploratory user study.

Chakrabarty, T.; Padmakumar, V.; Brahman, F.; and Muresan, S. 2024. Creativity support in the age of large language models: An empirical study involving professional writers. In *Creativity and Cognition*. New York, NY, USA: ACM.

Coenen, A.; Davis, L.; Ippolito, D.; Reif, E.; and Yuan, A. 2021. Wordcraft: a human-AI collaborative editor for story writing. *ArXiv*.

Creswell, J. W. 2012. *Qualitative inquiry and research design: Choosing among five approaches*. Thousand Oaks, CA: SAGE Publications, 3 edition.

Cursor. 2024. Cursor. https://www.cursor.com/. Accessed: 2024-9-11.

Dafoe, A.; Bachrach, Y.; Hadfield, G.; Horvitz, E.; Larson, K.; and Graepel, T. 2021. Cooperative AI: machines must learn to find common ground. *Nature* 593(7857):33–36.

Davis, N. M. 2013. Human-computer co-creativity: Blending human and computational creativity. In *Ninth Artificial Intelligence and Interactive Digital Entertainment Conference*. aaai.org.

Eno, B. 2024. AI's walking dog. https://www.bostonreview.net/forum_response/ais-walking-dog/. Accessed: 2024-12-30.

Essel, H. B.; Vlachopoulos, D.; Essuman, A. B.; and Amankwa, J. O. 2024. ChatGPT effects on cognitive skills of undergraduate students: Receiving instant responses from AI-based conversational large language models (LLMs). *Computers and Education: Artificial Intelligence* 6(100198):100198.

Gemini Team. 2024. Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context. *arXiv [cs.CL]*.

Gerlich, M. 2025. AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies (Basel)* 15(1):6.

Grigis, P., and De Angeli, A. 2024. Playwriting with large language models: Perceived features, interaction strategies and outcomes. In *Proceedings of the 2024 International Conference on Advanced Visual Interfaces*. New York, NY, USA: ACM.

Haase, J.; Djurica, D.; and Mendling, J. 2023. The art of inspiring creativity: Exploring the unique impact of AI-generated images. In *AMCIS 2023 Proceedings*.

Heaven, W. D. 2023. The inside story of how ChatGPT was built from the people who made it. *MIT Technology Review*.

Heersmink, R. 2024. Use of large language models might affect our cognitive skills. *Nat. Hum. Behav.* 1–2.

Horvitz, E. 1999. Principles of mixed-initiative user interfaces. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, CHI '99, 159–166. New York, NY, USA: Association for Computing Machinery.

Hwang, A. H.-C.; Liao, Q. V.; Blodgett, S. L.; Olteanu, A.; and Trischler, A. 2024. "it was 80% me, 20% AI": Seeking authenticity in co-writing with large language models. *arXiv [cs.HC]*.

Kantosalo, A.; Toivanen, J. M.; Xiao, P.; and Toivonen, H. 2016. From Isolation to Involvement: Adapting Machine Creativity Software to Support Human-Computer Co-Creation. 7.

Karimi, P.; Rezwana, J.; Siddiqui, S.; Maher, M. L.; and Dehbozorgi, N. 2020. Creative sketching partner: an analysis of human-AI co-creativity. In *Proceedings of the 25th International Conference on Intelligent User Interfaces*, IUI '20, 221–230. New York, NY, USA: Association for Computing Machinery.

Lee, M.; Gero, K. I.; Chung, J. J. Y.; Shum, S. B.; Raheja, V.; Shen, H.; Venugopalan, S.; Wambsganss, T.; Zhou, D.; Alghamdi, E. A.; August, T.; Bhat, A.;

Choksi, M. Z.; Dutta, S.; Guo, J. L. C.; Hoque, M. N.; Kim, Y.; Knight, S.; Neshaei, S. P.; Shibani, A.; Shrivastava, D.; Shroff, L.; Sergeyuk, A.; Stark, J.; Sterman, S.; Wang, S.; Bosselut, A.; Buschek, D.; Chang, J. C.; Chen, S.; Kreminski, M.; Park, J.; Pea, R.; Rho, E. H. R.; Shen, Z.; and Siangliulue, P. 2024. A design space for intelligent and interactive writing assistants. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, volume 7, 1–35. New York, NY, USA: ACM.

Lee, H. 2025. The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26-May 01, 2025, Yokohama, Japan* 1(1).

Masson, D.; Malacria, S.; Casiez, G.; and Vogel, D. 2024. DirectGPT: A direct manipulation interface to interact with large language models. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1–16. New York, NY, USA: ACM.

McGuire, J.; De Cremer, D.; and Van de Cruys, T. 2024. Establishing the importance of co-creation and self-efficacy in creative collaboration with artificial intelligence. *Sci. Rep.* 14(1):18525.

Muller, M.; Weisz, J. D.; and Geyer, W. 2020. Mixed initiative generative AI interfaces: An analytic framework for generative AI applications. *computationalcreativity.net*.

Norman, D. A. 2013. *The design of everyday things*. The MIT Press. London, England: MIT Press.

Ouyang, L.; Wu, J.; Jiang, X.; Almeida, D.; Wainwright, C. L.; Mishkin, P.; Zhang, C.; Agarwal, S.; Slama, K.; Ray, A.; Schulman, J.; Hilton, J.; Kelton, F.; Miller, L.; Simens, M.; Askell, A.; Welinder, P.; Christiano, P.; Leike, J.; and Lowe, R. 2022. Training language models to follow instructions with human feedback. *arXiv [cs.CL]*.

Palani, S., and Ramos, G. 2024. Evolving roles and workflows of creative practitioners in the age of generative AI. In *Creativity and Cognition*. New York, NY, USA: ACM.

Rezwana, J., and Maher, M. L. 2022a. Designing creative AI partners with COFI: A framework for modeling interaction in human-AI co-creative systems. *ACM Trans. Comput. Hum. Interact.*

Rezwana, J., and Maher, M. L. 2022b. Understanding user perceptions, collaborative experience and user engagement in different human-AI interaction designs for co-creative systems. In *Creativity and Cognition*. New York, NY, USA: ACM.

Shneiderman, B. 1997. Direct manipulation for comprehensible, predictable and controllable user interfaces. In *Proceedings of the 2nd international conference on Intelligent user interfaces - IUI '97*. New York, New York, USA: ACM Press.

Sun, Y.; Jang, E.; Ma, F.; and Wang, T. 2024. Generative AI in the wild: Prospects, challenges, and strategies. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, volume 18, 1–16. New York, NY, USA: ACM.

SXSW. 2023. OpenAI co-founder greg brockman on ChatGPT, DALL·E and the impact of generative AI — SXSW 2023.

Touvron, H.; Martin, L.; Stone, K.; Albert, P.; Almahairi, A.; Babaei, Y.; Bashlykov, N.; Batra, S.; Bhargava, P.; Bhosale, S.; Bikel, D.; Blecher, L.; Ferrer, C. C.; Chen, M.; Cucurull, G.; Esiobu, D.; Fernandes, J.; Fu, J.; Fu, W.; Fuller, B.; Gao, C.; Goswami, V.; Goyal, N.; Hartshorn, A.; Hosseini, S.; Hou, R.; Inan, H.; Kardas, M.; Kerkez, V.; Khabsa, M.; Kloumann, I.; Korenev, A.; Koura, P. S.; Lachaux, M.-A.; Lavril, T.; Lee, J.; Liskovich, D.; Lu, Y.; Mao, Y.; Martinet, X.; Mihaylov, T.; Mishra, P.; Molybog, I.; Nie, Y.; Poulton, A.; Reizenstein, J.; Rungta, R.; Saladi, K.; Schelten, A.; Silva, R.; Smith, E. M.; Subramanian, R.; Tan, X. E.; Tang, B.; Taylor, R.; Williams, A.; Kuan, J. X.; Xu, P.; Yan, Z.; Zarov, I.; Zhang, Y.; Fan, A.; Kambadur, M.; Narang, S.; Rodriguez, A.; Stojnic, R.; Edunov, S.; and Scialom, T. 2023. Llama 2: Open foundation and fine-tuned chat models. *arXiv [cs.CL]*.

Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, ; and Polosukhin, I. 2017. Attention is all you need. *Adv. Neural Inf. Process. Syst.* 30.

Wan, Q.; Hu, S.; Zhang, Y.; Wang, P.; Wen, B.; and Lu, Z. 2023. "it felt like having a second mind": Investigating human-AI co-creativity in prewriting with large language models. *arXiv [cs.HC]*.

Yannakakis, G. N.; Liapis, A.; and Alexopoulos, C. 2014. Mixed-initiative co-creativity.

Yuan, A.; Coenen, A.; Reif, E.; and Ippolito, D. 2022. Wordcraft: Story writing with large language models. In *Proceedings of the 27th International Conference on Intelligent User Interfaces*, IUI '22, 841–852. New York, NY, USA: Association for Computing Machinery.