

Ai.Illude: Encouraging Rewriting AI-Generated Text to Support Creative Expression

David Zhou
University of Illinois Urbana-Champaign
Urbana, USA
david23@illinois.edu

Sarah Sterman
University of Illinois Urbana-Champaign
Urbana, USA
ssterman@illinois.edu

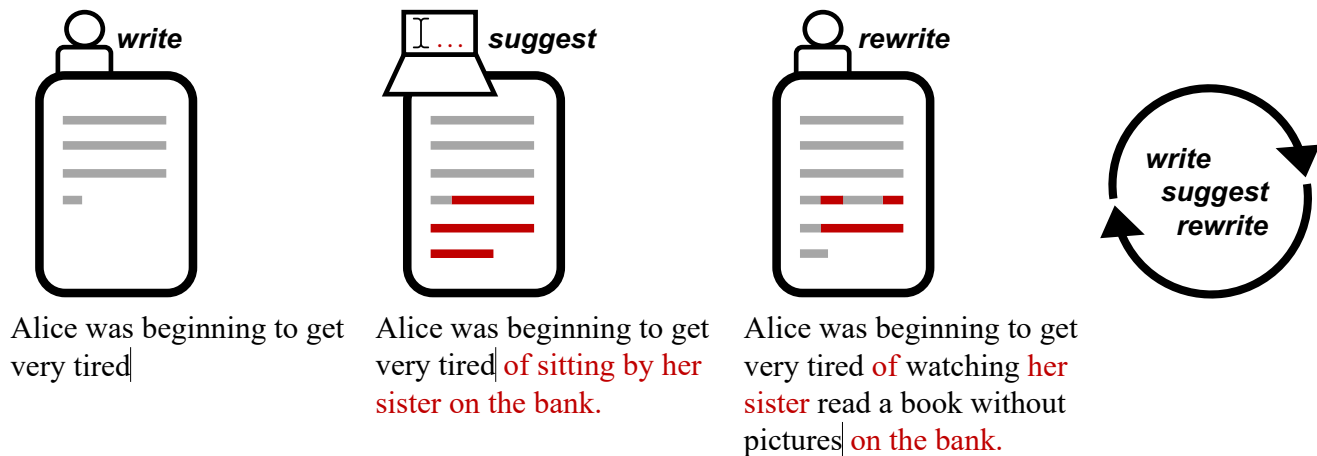


Figure 1: Rewriting is an important stage of creative writing, influencing self-expression, creative goals, and final output. This paper studies how writers adapt AI suggestions and evaluates the effect of rewriting AI-generated suggestions on writing process. Figure text: adapted from Alice in Wonderland (Lewis Carroll).

ABSTRACT

In each step of the creative writing process, writers must grapple with their creative goals and individual perspectives. This process affects the writer’s sense of authenticity and their engagement with the written output. Fluent text generation by AIs risks undermining the reflective loop of rewriting. We hypothesize that deliberately generating imperfect *intermediate text* can encourage rewriting and prompt higher level decision making. Using logs from 27 writing sessions using a text generation AI, we characterize how writers adapt and rewrite AI suggestions, and show that intermediate suggestions significantly motivate and increase rewriting. We discuss the implications of this finding, and future steps for investigating how to leverage intermediate text in AI writing support tools to support ownership over creative expression.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**; • **Applied computing** → **Arts and humanities**.

KEYWORDS

intelligent writing assistant, creative writing, text generation, editing, revising, AI writing

ACM Reference Format:

David Zhou and Sarah Sterman. 2024. *Ai.Illude: Encouraging Rewriting AI-Generated Text to Support Creative Expression*. In *Creativity and Cognition (C&C '24)*, June 23–26, 2024, Chicago, IL, USA. ACM, New York, NY, USA, 14 pages. <https://doi.org/10.1145/3635636.3656187>

1 INTRODUCTION

Creative writing is a form of self-dialogue [15, 32]. Through the process of writing, writers reflect on not only language, but also aesthetics, imagination, and identity. Advances in language models have led to waves of new intelligent writing assistants designed to make writing easier by generating fluent text [29, 36]. While these tools can speed up the generation of content or help writers overcome writer’s block [12, 49, 50], it is important to consider how computational tools, particularly ones that take over self-expression, affect personal writing processes and goals.

One significant concern is that by removing the writer from the process of making difficult choices about ideas and words, such tools may also reduce the space for deep thinking that occurs through struggle. Feelings of authenticity and ownership are associated with the struggle to create inventive ideas [2, 39] and the investment of effort, attention, and time into the textual artifact [16]. In recent work, the need for user control and ownership have emerged as potential design guidelines for intelligent writing tools [3, 27].

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
C&C '24, June 23–26, 2024, Chicago, IL, USA
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0485-7/24/06.
<https://doi.org/10.1145/3635636.3656187>

Therefore a key question arises for the design of intelligent writing tools: what aspects of intelligent writing assistance increase or undermine writers' control over the work, and how do we balance the benefits of intelligent writing assistants, such as improvements in writing efficiency or democratization of writing skills, with the need for creative writers to have space for critical reflection? We hypothesize that rewriting can be an important dimension for understanding how AI text generators influence feelings of control. As a first step, this paper looks at rewriting AI-generated text as a way to support control over writing process and output.

Rewriting is an essential part of the creative writing process; during rewriting, language is refined, removed, or inserted in order to change, clarify, or expand the ideas of the text. In doing so, the writer grapples both with ideas and the effective expression of those ideas through words and structure. Prior work has studied the relationship between rewriting AI-generated text and psychological ownership in limited forms [35]. However, significant questions remain around the effect of rewriting on writer's process, understanding why and how writers rewrite AI-generated text, and defining rewriting in the context of AI generation. In this work, we address the following questions: *How can AI suggestions be designed to encourage the rewriting of AI-generated text?* and *How does the rewriting of AI-generated text affect the writer's personal writing process?*

To explore the rewriting of AI-generated text, this paper presents a study of 27 creative writers using a custom text editor that provides two types of AI-generated suggestions: first, a *fluent continuation* to follow the user's text; second, an *intermediate suggestion* consisting of multiple fragmentary ideas. For the first, we request a completion of the user's text without additional prompts; the second is created by concatenating four potential ideas about plot and setting into a single continuation. Fluent continuations are designed to be grammatical and coherent to context. Intermediate suggestions are meant to provide relevant content that cannot be directly incorporated into the text. Instead, suggestions must deliberately be altered to become grammatical and coherent. We hypothesize that shifting the focus of generative writing assistants away from creating perfect output to creating intermediate text that encourages rewriting will support the personal and reflective aspects of creative writing.

We use three metrics to capture different aspects of rewriting behaviors: *remaining AI text*, *sentence embedding similarity*, and *number of user edits*. We use these metrics, along with qualitative analysis, to evaluate rewriting behaviors and writing engagement, as well as characterizing the difference between these two completion paradigms. We find that writers tend to keep less of intermediate suggestions compared to fluent continuations. Despite being perceived as more flawed, we find that writers are able to adopt intermediate suggestions more diversely; rather than simply for writing continuation, writers use these suggestions to help brainstorm ideas, even if none of the original suggestion remains in the final story.

This paper contributes:

- (1) The concept of *intermediate text* and an implementation of an *intermediate suggestion*, a paradigm for engaging creative writers in editing AI-generated text.

- (2) Results from a 27-person study of writing with *intermediate suggestions*:
 - characterizing rewriting behaviors on AI-generated text in creative writing
 - showing intermediate suggestions increase rewriting
 - showing the effects of suggestion design on writing process: fluent continuations support writing ease, while intermediate suggestions support ideation and reflection.
- (3) Discussion of designing generative writing assistants that center writers' process and ownership.

2 RELATED WORK

2.1 Creative Writing and Rewriting

Creative writing is deeply personal, open-ended, and connected to self-expression, making the process of how we write an important subject of study. There are no unambiguous definitions of acceptable outcomes for creative tasks, nor are there obvious and predetermined paths to creative goals [1]. Instead, the process of seeking a creative goal can be valuable in itself and essential to the character of the final output. Writing is an act of *making* meaning, not finding it [22], and the act of writing can reciprocally shape the author as well [39]. When we consider that the tools and environments we use shape our process, not only in how we take actions but even in what actions we can imagine or are available to take [17, 23], it becomes essential to ask how our writing support tools affect writers' process of expression.

Rewriting is a vital step in the writing process. Rewriting is not only the act of writing again [20]; it is intertwined with the overall effort to *make it right* – to move from word to phrase to sentence to paragraph then back to words, to see the need for variety and balance, firmer structure, and more appropriate form [38]. Writing, as a process of meaning-making and discovery, can be characterized as a reactive process between writer and textual artifact [23, 25]. This process involves reading and rereading, as well as writing and rewriting [38], a form of hands-on thinking [9]. In his book *Writing Without Teachers*, Elbow describes writing as words gradually changing and evolving [21]. Authors of classic literature Leo Tolstoy and Roald Dahl describe rewriting as a reflective practice, one where writing goals develop and completely change over time [38]; Ernest Hemingway rewrote the ending to his first bestseller *Farewell to Arms* thirty-nine times [42]. Since rewriting is a reciprocal process, it is important to understand what factors influence rewriting and what factors motivate it in the author.

Most writing now is done on computers with some form of word processor. Since tools fundamentally shape process, we must consider how they affect rewriting, and therefore meaning-making and self-expression. The effects of word processors on writing process have been studied for decades; for example, writing with a word processor has led to more local revisions (word, phrase, sentence level) compared to writing by hand [14, 19]; for novice writers, these changes tended to be lexical [51] while experienced writers tended to make global changes [31]. The effects of AI text generation tools are still uncertain, especially as the capability of AI systems to output fluent text has recently increased dramatically. In this paper, we explore how two variants of text generator outputs

affect rewriting behaviors and how writers perceive AI writing in context of their personal writing processes.

2.2 AI Writing Support Tools

Writing support tools have been a mainstay of creativity support research and development in HCI since the early days of computing [40]. Recent works have proposed specialized tools to support greater control over expression, such as word finding and writing style [24, 26]. Studies have shown that writers have diverse expectations when integrating computational assistance into their existing creative writing process [5, 33]. Recent studies have also explored the need for self-reflection, such as by watching real-time writing replays [10], and conversing with a character development tool to explore deeper characterization [47].

In recent years, a rapid increase in fluency and capability of large language models has driven a surge of AI-supported writing tools. Here we discuss three types of outputs for generative writing tool design: *fluent multipurpose generation*; *task-specific targeted suggestions*; and *resources for interpretation*. In the next section, we discuss the design goal of encouraging *rewriting* in relationship to AI features.

Fluent multipurpose generation: Large language models now have the capability to generate contextually relevant and fluent text. One genre of AI writing support tools use this ability to generate completions or suggestions that are inserted directly into the user's text [12, 29, 35, 45, 46, 53]. These suggestions are designed to be relevant, grammatically correct and potentially included into the final output without modification. Such tools offer completions during the writing process, and are motivated by goals like helping the user overcome writer's block, provide new ideas, or speed up the writing process [12, 49, 50].

Task-specific targeted suggestions: A second genre of AI writing support tools focuses on supporting one particular type of writing task, providing targeted suggestions rather than general text completion. For example, Metaphoria [27] helps writers create metaphors by suggesting connections to ideas and words. Inkwell provides stylistic variations of texts [24] to assist style goals in poetry. A proposed character backstory generation system [13] helps game designers create convincing characters in immersive environments. By targeting specific writing tasks, these tools intervene only in a limited part of the writing process.

Resources for interpretation: Writing support tools can also provide resources for interpretation as their text generation goal. Rather than providing text meant to be included in the output, these tools provide annotations, feedback, or inspirational media. For example, multimodal generation provides writers non-text suggestions that force the writer to evaluate and interpret provided feedback [49]. Hai et al. propose providing writers with paragraph summaries as margin annotations to support revision [18]. These AI generations are designed to motivate writing processes while leaving greater control and responsibility over choices about the text itself to the user.

In each of these tools, the usefulness and perceived quality of its suggestions are often limited by the technical capabilities of the underlying language model. Despite these shortcomings, AI

text generators do not necessarily require ideal coherency to support creativity processes. Literary cut-up techniques, a method of creating new texts by combining the fragments of existing texts, employ chance to generate novelty [30, 44]. The stochasticity of language models may also function as a source for inspiration — to inspire writers to make cognitive leaps between elements of semantic relevance in order to further their creative goals [49].

2.3 Motivating Rewriting

Flower and Hayes identify the task environment as an influential external factor of the cognitive writing process [23]. AI writing support features alter the task environment and affect both engagement with writing and expression of personal perspective. Generative writing assistants that directly insert text into the writing change the task environment and have the potential to negatively affect feelings of engagement, particularly if the provided suggestions are very good [27]. Jakesch et al. show that suggestions provided by an opinionated language model alter the opinions writers express in their work [34]. To design AI writing assistants that can help with writing tasks without undermining the writer's exploration of meaning and expression, we must consider how to increase engagement and provide greater control to the writer over their process. This paper takes up *rewriting* as one design direction to support ownership and expression.

Rewriting is an integral piece of the writing process, an opportunity to find new and surprising ideas and learn from one's own expression [37]. In reflective practice, the practitioner changes their understanding of the problem space and updates the actions they take through the process of doing [48]; for creative writers, rewriting can be a powerful source of reflection-in-action. The connection between rewriting and the construction of goals and self-expression underlies our decision to investigate rewriting as a potential mediator of psychological ownership.

Lee et al. performed a preliminary analysis of rewriting and ownership by looking at rewriting behaviors across the entire text [35]. They found a weak, insignificant relation between the number of edits, approximated by the number of selection and delete events, across the entire text and ownership. However, they were not able to investigate the rewriting of AI suggestions specifically. We seek to deepen our understanding of the relationship between rewriting and ownership by evaluating rewriting specifically done on AI-generated text with a more extensive survey assessing control over self-expression.

We also seek to understand how to create AI-generated text that encourages rewriting. We build on the motivation of systems that generate resources for interpretation (Section 2.2): providing material for the user to interpret, rather than solutions to incorporate. We design a system to generate intermediate suggestions, which must be rewritten by the user before it can be coherently incorporated into the story. While older AI tools tend to focus on the benefit of serendipity to the writing process and are limited by the coherence of the models [8, 24, 46], the recent increase in the capability of language models to generate fluent and relevant text enables us to design deliberate alterations to coherency.

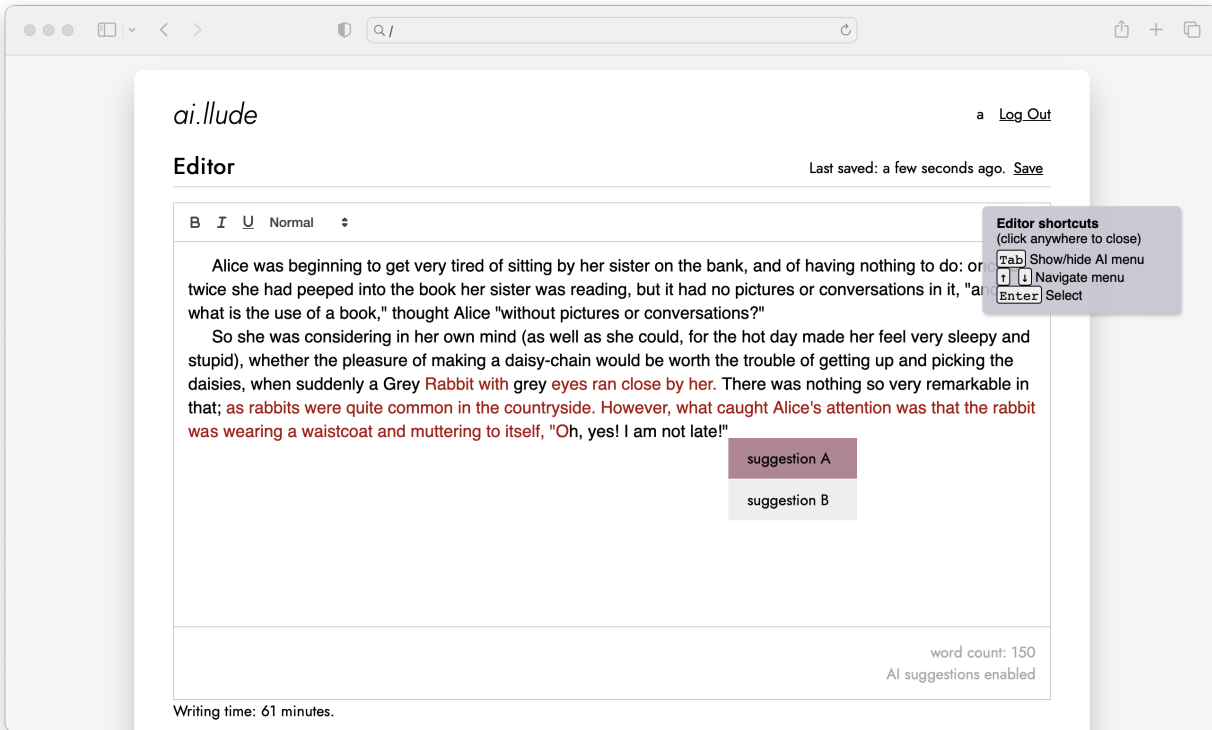


Figure 2: Interface of *ai.llude*. The AI menu opens when the tab key is pressed and allows writers to request suggestions. User text is displayed in black and AI generated text in purple. The word count and writing time are displayed at the bottom of the editor. The editor provides basic styling controls such as bold, italic, and underline. Figure text: Alice in Wonderland (Lewis Carroll).

3 DESIGN OF AI.LLUDE

To study rewriting behaviors, we built a system for data collection to explore how we might encourage rewriting. *Ai.llude* is an instrumented text editor with AI support designed to encourage rewriting. It is designed to simulate a realistic writing environment and provides a streamlined interface for working with AI suggestions, as well as extensive logging capabilities.

3.1 Design Goals

Past work has identified **interactivity** and **control** in AI writing assistants as design characteristics that affect perceptions of engagement with the assistant and **ownership** over the writing artifact [3, 27]. These principles are not exclusive to digital writing assistants. For human-to-human design collaboration, defining scope and encouraging conversation make creation and collaboration more clear and engaging [28], while the emphasis of voice and identity in writing contexts in general provides the writer with greater agency and expression [4].

To address the need for greater interactivity and control, we explore the paradigm of generating intermediate text. In contrast to [18, 49], which generate intermediate interpretations that exist outside of the created artifact (e.g. images, margin summaries),

intermediate text consists of suggestions that are part of the creative artifact itself but do not remain as-is in the final outcome. The concept of intermediate text was inspired by the initial writing stages where writers jot down ideas before they iteratively revise their writing. As the process of writing not only changes the textual artifact but also writing goals, the final writing may be drastically different than the imagined outcome that was desired in the beginning.

3.2 System Overview

Ai.llude is a Flask-based web application (Figure 2). It contains a customized Quill.js text editor [11], where writers can compose stories and request AI suggestions, which are continuations of the text in the editor. Interactions with the editor and with the AI suggestions interface are comprehensively logged; complete writing sessions can be reconstructed.

To make the AI user-interface more streamlined for text editing, the editor displays a drop-down menu at the caret location when the tab key is pressed, inspired by CoAuthor [35]. The drop-down menu displays two possible AI suggestions: **Suggestion A** and **Suggestion B**, corresponding to *fluent continuation* and *intermediate suggestion*. These suggestion types are named generically in

the interface to prevent nomenclature bias. Figure 2 shows the drop-down menu.

All suggestions are requested and received using the OpenAI Chat completions API and generated with GPT-3.5 Turbo. Each request contains the story from the beginning to the current caret location. The returned suggestion is inserted after the caret in order to let AI suggestions be used anywhere in the story. If the provided story text is too short (less than 100 characters), the AI suggestions are disabled to reduce the likelihood of an irrelevant suggestion. Generated text is inserted in purple to visually distinguish human and AI text. When the user types, the user-input characters are presented in black.

Suggestion A provides AI generated text that is designed to be a *fluent continuation*. The text up to the caret is provided to the model for completion; no additional system or user prompts are provided, and no parameters are changed from the default (e.g. temperature = 1), so a fluent continuation mimics the style and tone of the story. As this suggestion provides the statistically optimal continuation of the text, we imagine that this type of suggestion has relatively **low interactivity**; i.e., less potential for revision and less room for critical engagement. We hypothesize that this type of AI generation does little to encourage rewriting. To prevent suggestions from overwhelming the writer and story, Suggestion A suggestions are terminated after a newline (signifying a new paragraph) or after 60 tokens are generated, whichever comes first.

Suggestion B inserts a sequence of unrelated ideas concatenated into a single string, designed not to be used as part of the written artifact as-is. It is designed to be a prototype of *intermediate text*. Specifically, this suggestion queries the Chat completions API independently four times with prompts designed to return the following types of continuations:

- (1) **Plot Beat**. Suggests a shift in narrative; e.g., new events, actions, emotional turns, realizations, etc. This is designed to suggest a new plot element that should be further expanded if used.
- (2) **Setting Detail**. Adds a new detail about the setting, such as new sensory information or a new setting element; e.g., a flowering tree if the setting is a courtyard. This is designed to expand the writer’s imagining of the scene and should be expanded if used.
- (3) **Benign Tangent**. Suggests a benign circumstantial event occurring in the background; e.g., a squirrel scurrying across the grass in a fishing scene. This is designed to expand the writer’s imagining of immersion and should be expanded if used.
- (4) **Whimsical Storyteller**. Provides a continuation of the story in a light-hearted and joking tone. This is designed to be a stylized continuation of the story that should be rewritten into a more similar style to the story if used.

Each individual query is instructed to cap itself at 15 words, with a 30 token hard limit, in order to approximate the length of Suggestion A. Table 1 shows examples of both types of suggestions.

Suggestion B was developed through multiple iterations, with the initial goal of requiring editing to be grammatical. Early versions included removing punctuation and capitalization to create lengthy, ungrammatical run-on sentences inspired by stream of

consciousness writing. However, pilot testing showed that this design impeded usability because it required too much effort to read and was too difficult to understand. To improve ease of use and relevance to creative writing, we decided to append multiple unrelated continuations together. This was designed to be easy to read, but still ill-formed to encourage rewriting. **Plot Beat** and **Benign Tangent** are designed to assist with the development of the story’s direction, while **Setting Detail** and **Whimsical Storyteller** are designed to assist with crafting prose.

In general, we design intermediate suggestions to have **high interactivity** and to encourage **greater control** over creative decisions. First we imagine a selection process to take place: one component might be chosen and the rest might be discarded. Secondly, we imagine that due to the suggestions being underdeveloped or improperly stylized, any desired components will be further rewritten. We hypothesize that combining selection with rewriting will make the suggestion more interactive.

4 METHODS

To evaluate the effects of AI writing on rewriting, we asked creative writers to each write a short story using *ai.llude*. We received IRB approval from our university in order to evaluate the system on human participants. Study risks were stated in the consent form: (1) privacy concerns of using an online system and (2) the potential risk to receive harmful AI-generated content. To help mitigate (1), participants were informed that they can prevent sending editor interaction data to the server by refreshing the editor without explicitly saving, and an enterprise OpenAI account was used to have complete ownership of all data. To help mitigate (2), a language filter was used to prevent harmful language returned by the Chat completions API from being displayed in the text editor. If a completion included a blacklisted word, it would not be shown.

4.1 Participant Recruitment

People who self-identified as creative writers and consented to the study were recruited from university mailing lists, fliers posted in public locations (e.g. a local book store), and by word-of-mouth. We screened study applicants using their reported creative writing history: we recruited participants who had previously engaged in creative writing for personal expression or enjoyment. Study applicants who did not previously engage in creative writing or only wrote personal essays (e.g. college application essays) were not included. Thirty-two participants were recruited for the study. Five participants did not write for the full session, or only pressed the AI suggestion button and did not write any text; these sessions were excluded from the data analysis. This left 27 participants who completed the entire study.

Of these, 11 identified as male, 14 identified as female, 1 self-described, and 1 did not disclose. The average age was 22.7 years (min: 18, max: 32). Nine participants self-identified as a beginner writer; 14 self-identified as an intermediate writer, and 4 self-identified as an expert writer. The three most commonly self-reported types of writing were short fiction (17), journals (12), and poetry (11). Fifteen participants had earned a bachelor’s degree, and 12 were currently undergraduate students. While we did not restrict recruitment, most participants were college students or recently

<p>User Text Alice was beginning to get very tired of sitting by her sister on the bank, and of having nothing to do: once or twice she had peeped into the book her sister was reading, but it had no pictures or conversations in it, "and what is the use of a book," thought Alice "without pictures or conversations?"</p>	
<p>Fluent Continuation So, she decided to explore the riverbank. As she wandered, she stumbled upon a peculiar looking rabbit hole hidden behind some bushes. Curiosity piqued, Alice peered into the dark abyss, wondering where it might lead. With a deep breath, she made a bold decision to follow the white rabbit, diving...</p>	<p>Intermediate Suggestion Suddenly, a peculiar-looking creature appeared with a book that could bring pictures to life... Just as Alice was about to resign herself to the monotony... Suddenly, a squirrel scurried up to Alice and sat on her lap... The bank where Alice and her sister sat was lush with wildflowers of every color imaginable...</p>

Table 1: Two types of text completions can be requested. Top: user text. Bottom left: a potential fluent continuation. Bottom right: a potential intermediate suggestion. Figure user text: Alice in Wonderland (Lewis Carroll).

graduated from college, with an average of 6.3 years of writing experience. Due to the limited age range of the sample set, and geographic similarity, our results are likely influenced by demographic characteristics.

In the pre-survey, participants mostly felt that AI text generators were *good at generating fluent text* (agree: 21, neutral: 5, disagree: 1) and at *providing useful responses* (agree: 19, neutral: 8), but participants were divided on whether AI text generators were *good at providing new ideas* (agree: 7, neutral: 13, disagree: 7). Study participants were compensated \$10 USD.

4.2 Experimental Design

4.2.1 Writing Environment. To create a realistic and engaging writing environment, writers were asked to imagine a fictitious scenario of writing a short story to be submitted to a literary column. Short story writing is a common task for studying writing assistants [8, 10, 12, 35, 45, 46, 49]; it is a task familiar to participants and relatively constrained. Participants were provided intentionally vague descriptions of both fluent and intermediate suggestions to encourage experimentation and completed a short tutorial session. In the tutorial, participants were encouraged to try both suggestion types on a provided story excerpt to acclimate to the writing environment and see examples of potential AI output. To help provide writing inspiration, three optional writing prompts were provided. Writers were instructed to write for at least 50 minutes. The AI suggestion menu was disabled for the first 15 minutes of writing, or until 150 words were written, to encourage writers to develop initial ideas and story-writing plans without AI influence. To reduce order effects, the tab menu presented Suggestion A on top to half the participants, and Suggestion B on top to the other half. Writing prompts and user instructions are provided in the supplemental materials.

4.2.2 Writing Evaluation Survey. At the end of the writing session, participants filled out an exit survey in which they reflected on their experiences using the AI suggestions. To evaluate the effects of fluent and intermediate suggestions, the exit survey probed overall feelings of satisfaction (e.g. "are you proud of your story") and self-concept ("e.g. does your writing sound like you") [2, 52]. By giving

the survey at the end of the writing session, these questions were designed to gauge attitudes towards the overall writing outcome and to capture the overall process of using and integrating AI assistance.

Participants were also asked to reflect on each suggestion type by considering self-concept ("The suggestion sounded like something I could have written") and responsibility ("I felt the need to rewrite what the suggestions gave me") [2, 52]. This subset of questions was designed to assess whether suggestions were aligned with the writer's personal intentions and writing goals. Finally, participants were asked to reflect on how they used, evaluated, and integrated both type of suggestions as open-ended responses. The survey was informed by prior literature on psychological ownership and control [41, 52], which we chose as a way to assess self-expression. The full list of survey questions can be found in the supplemental materials.

4.2.3 Rewrite Metrics. We define rewriting as a series of actions that changes an original span of text to the version of that text in the final document. In this work, we are interested in the changes that people make when they rewrite AI suggestions. Therefore, we investigate rewriting by looking at the differences between the original AI suggestion and the final rewritten version. Here, we describe both how we operationalize initial and final spans of text, and the metrics used to characterize the differences. We define the "original suggestion" as the span of text initially appended by the AI to the editor. We define the final "rewritten version" as the span of text in the final draft that begins at the first remaining AI character from the original suggestion, and ends at the final remaining AI character, i.e. the suggestion after new text insertions and deletions. See Table 3 for an example. To evaluate the extent of rewriting, we measure the difference between the original AI suggestion and its final rewritten version using three metrics:

- (1) Words Remaining.** This is the percentage of AI text that was kept from the original suggestion in the final rewritten version; specifically, the number of AI words in the final rewritten version divided by the number of AI words in the original suggestion.

<p>"I mean, I think that maybe we should just take a break and see what it might mean. Maybe there's something we're missing, you know?" Theo suggests, his voice tinged with uncertainty. "Maybe we're both holding each other back from something."</p> <p>Lucy's heart sinks as Theo's words hang in the air. A break? It feels like the ground beneath her has given way. She has to look out the window so that he doesn't see the tears accumulating in her eyes. The city seems to be in constant motion while her world has just stopped.</p>	<p><i>Fluent continuation.</i></p> <p><i>Intermediate suggestion, edited to keep only a portion of the suggestion.</i></p> <p><i>Fluent continuation, rewritten to incorporate multiple segments of user text.</i></p>
--	--

Table 2: An excerpt of participant W18's story, showing how they integrated both types of AI suggestions. User text is presented in black. Text generated by Suggestion A is in blue, and Suggestion B in orange. The portion of the intermediate suggestion kept in the story was the "whimsical" continuation. Note that in the *ai.llude* interface, all AI text is presented in a single color.

- (2) **Sentence Embedding Distance.** This measures semantic similarity between the original suggestion and the rewritten version using cosine similarity of the two computed vectors calculated by [43].
- (3) **Number of Edits.** This counts the number of text-insert and text-delete operations occurring within each suggestion. The number of user edits is a measure of the effort it took to rewrite the original suggestion. Counting text operations is important for measuring rewriting since writers might make repeated changes that might not be captured by comparing the original and final states, such as typing and backspacing or undoing and redoing.

We calculate these metrics for all AI suggestions, whether A or B, and compare the results for these metrics between suggestion types.

4.2.4 Qualitative Analysis. To contextualize the quantitative results with how and why participants rewrote AI suggestions, we conducted a reflexive thematic analysis [6] on the open-ended responses of the exit survey (Section 4.2.2). Because rewriting AI text is an underexplored area, inductive coding allows us to independently build up our understanding of rewriting behaviors from the data. Both authors coded a subset of survey responses, then discussed their codes to share interpretations and reach consensus. Then the first author open-coded the remaining responses and grouped the codes into potential themes. That author then checked each code and theme again to refine groupings and merge similar themes. The themes were discussed and refined again by both authors.

5 RESULTS

We analyze the data collected from our study in three ways. First, we characterize overall writing and rewriting behaviors of participants when using *ai.llude*, finding two predominant themes of AI text rewriting: *stylistic consistency* (Section 5.1.1), and *selection process* (Section 5.1.2). Then, we compare writing and rewriting behaviors between fluent continuations and intermediate suggestions, and show that the design of intermediate suggestions significantly

increases rewriting (Sections 5.2.1 & 5.2.2). We address the relationship between suggestion design and writing process, finding that fluent continuations make writing easier and intermediate suggestions support reflective thinking (Section 5.3). Finally, we report how suggestions were adapted in sometimes unintended ways, an illustration of creative misuse (Section 5.4).

5.1 Characterizing Writing Process and Rewriting Behaviors

Stories written with *ai.llude* were on average 1,205 words long, equivalent to approximately 2 pages of 12pt single-spaced text (min: 218 words, max: 2,969 words). Stories contained on average 370 AI generated words, i.e. 31% of total text (min: 0 words, max: 2,197 words). The average writing session lasted for 62.7 minutes (min: 46.0 minutes, max: 115.3 minutes). Table 2 shows an excerpt of a participant-written story, demonstrating how a writer might integrate suggestions into their story and how the final text alternates between AI-generated and user-written content.

Figure 3 depicts timelines of each writing session, indicating when the user triggers AI suggestions and when the user types. Some writers traded off with AI suggestions and requested them sparingly, with lengthy periods of their own writing in-between each suggestion (W2, W7, W16). Others used suggestions more frequently, sometimes in bursts, for example deleting unhelpful suggestions until landing on a useful one (W3, W19, W25). Writers varied in their usages of both suggestion types: one writer used Suggestion B once, then stuck with Suggestion A for the rest of the writing session (W24); one used Suggestion B more often (W19). Others tended to switch between both types equally.

After requesting AI suggestions, writers take one of three actions: 1) they delete the entire suggestion, 2) they leave the entire suggestion as-is, or 3) they rewrite part or all of the suggestion, including deleting portions of the suggestion. Suggestions with one segment contain no internal user segments; e.g., text left as-is (i.e. one AI segment and no user segments). Writers would also oftentimes delete large portions of generated text and use a small but contiguous piece, which is also one AI segment and no user segments. Fewer segments are more common; 92% of suggestions that are kept in the text include five or fewer segments. Table 3 shows

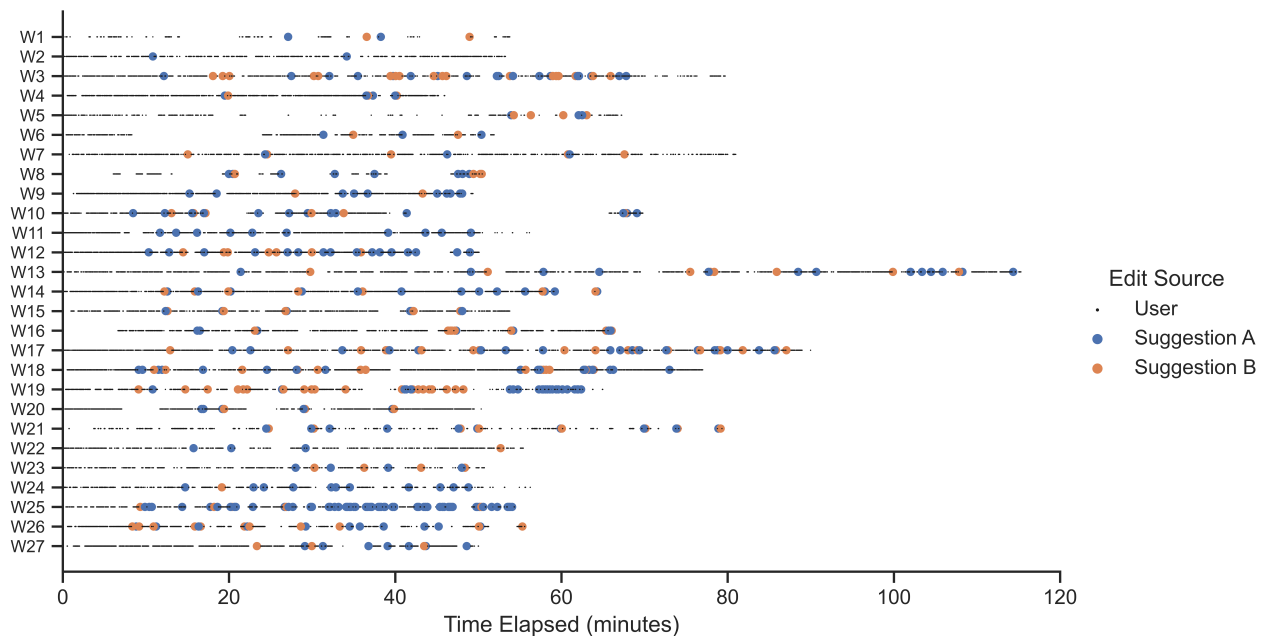


Figure 3: Writing Flows. Dots indicate activity in the text editor, either of the user typing or the user requesting an AI suggestion. The suggestions were disabled for the first 15 minutes of the session, or until 150 words had been written.

an initial AI continuation provided by a fluent continuation (surrounding user text provided for context), which was then rewritten by the user. This rewriting resulted in nine alternating segments: a large insertion describing the narrator’s backstory, followed by small grammatical edits to set the suggested setting description to past tense.

5.1.1 Stylistic consistency. AI suggestions were helpful for introducing novelty, inducing effective surprise in the writer. To integrate suggestions into their stories, writers would keep the core concept and rewrite the suggestion to make it stylistically fit with the writer’s intentions. This type of rewriting included adapting the text to the writer’s own voice, changing the style of text to align with the writing subject, and ensuring consistency with surrounding style:

I merely changed the suggestion to make it sound a bit more like what I would write. (W4)

If I liked the suggestion I would make it stylistically fit in with the prompt. (W1)

I usually reworked any of the suggestions by modifying the words used to fit the formality of the passage or by reworking its flow so that it sounded better. (W10)

Conversely, some writers also used AI suggestions to seek wording and style, discarding the idea or direction contained in the suggestion:

If I didn’t like the ideas I would see if I like the style of writing could fit. (W1)

For instance, if I needed the name of something, I would use Suggestion A and regenerate it until it produced a name I liked. (W18)

Generally, the value of a suggestion depended on its **novelty** (whether it contained new ideas, directions, alternatives, or interesting plot elements) and its **stylistic coherence** (whether it aligned with writing style, tone, voice, flow, and creative intentions). Suggestions did not need to contain all of these elements to remain in the story or be useful to the writer. Ones that contained novel ideas, but with undesired style, were integrated and rewritten for stylistic consistency; others with relevant style, but uninteresting ideas or direction, were used to assist with wording. These insights build off of Singh et al.’s findings that in some cases, writers were not hindered by less semantically relevant suggestions or less linguistically coherent sentences [49].

5.1.2 Selection process. AI suggestions, particularly the intermediate suggestions provided by Suggestion B, tended to contain elements that were not all useful for the writer. This would result in a selection process: writers used suggestions to consider smaller phrases and alternatives, and chose the pieces that they wanted and discarded the rest. The selection process could, but did not necessarily, co-occur with integrating style; the selection process also helped writers consider alternatives and potential directions for their story:

I picked and chose those which could help me better my story. (W4)

I did use fragments of some suggestions (I think mostly Suggestion A) to fill in phrases I needed but couldn’t come up with. (W7)

Initial Suggestion	Rewritten Suggestion
<p>... because that was some much needed confidence boost for the nerve-wracked me. I take a deep breath and straighten my tie, making sure every detail is perfect. The Rabbit Diner is a quaint little place, known for its cozy atmosphere and delicious comfort food. As I step inside, the warm aroma of coffee and pancakes envelops me, instantly calming my nerves and bringing my attention to the more important matter at hand, my blind date.</p>	<p>... because that was some much needed confidence boost for the nerve-wracked me. I take a deep breath and straighten my tie, making sure every detail is perfect. I almost get a heart attack with a almost perfectly timed beep from my pager. Yeah you heard (or read) that right. A beep from my PAGER. Call me a old soul or whatever but nothing screams security like good ol'fashioned technology. I dont have space in my head for tiktok or trends. Throw any language or subject at me, and I can bore you to death with knowledge about it. You want me to dance like a toy at the whims of the current trends, both literally and figuratively? Count me out. "Red rabbit at yellow gates at full moon ", the pager states. That immediately distracts me from my nerves. Apparently, my life-long nemesis(long story short, we hate each other and work for opposite sides) is a lady with red hair and she has a meeting with some important clients at the gates of the Pakistan museum at noon. How did they even discover all of this information when I was never able to even catch a glimpse at their face for all these years?</p> <p>The Rabbit Diner was a quaint little place, known for its cozy atmosphere and delicious comfort food and as I stepped inside, the warm aroma of coffee and pancakes envelops me, instantly calming my nerves and bringing my attention to the more important matter at hand, my blind date.</p>

Table 3: Two versions of an excerpt of participant W9’s story. Left: An AI suggestion inserted into the user text. Right: The same excerpt after the user rewrote the suggestion, including a large insertion of character backstory, and small grammatical edits. Text from the original suggestion is presented in blue. Surrounding user-written text provided for context.

"B" often had one or two usable (given the context) possibilities, but went into a lot of irrelevant and generic detail. (W19)

The selection process was not limited to immediate integration of the suggestions. Some writers also selected pieces of suggestions for *distant use*:

Suggestion B was usually a bit more messy and incoherent, so I would split up Suggestion B and maybe use it different places. (W6)

If suggestion B had a part that I liked but didn't fit then I would make a note of it to use later in the story. (W10)

Similar to rewriting for stylistic consistency, the value of a suggestion depended on novelty and stylistic coherence, and did not need both to be fragmented and selected (choosing fragments) by the writer. The selection process enabled writers to assess and integrate suggestions **recursively**; if an entire suggestion was unusable as-is, there could still be value in integrating a smaller, but more useful fragment, while deleting the rest. Thus, the selection process surfaced as a subprocess of integrating AI suggestions.

5.2 Assessing Usage of Fluent Continuations vs Intermediate Suggestions

5.2.1 *Intermediate suggestions increase rewriting.* Writers tended to keep Suggestion A in its entirety, rewrite it to some extent, or reject it, whereas very few Suggestion B suggestions were kept in their entirety. Figures 4a and 4b compare original and rewritten suggestions according to remaining words (0: no words kept; 1: all words

kept) and sentence embedding similarity (0: not semantically similar; 1: very semantically similar). **Writers kept significantly less of intermediate suggestions than fluent suggestions, both on page and semantically:** fluent suggestions had an average of 60% remaining text and intermediate suggestions had an average of 16% ($p = 0.00, t = 11.51$). Using sentence embedding similarity, fluent continuation had an average of 0.63 and intermediate suggestions had an average of 0.28 ($p = 0.00, t = 8.70$).

The average text operation count (number of text-insert and text-delete events) between the provided suggestion and the final text was significantly different for fluent continuations and intermediate suggestions as well (*fluent continuation*: 444, *intermediate suggestion*: 795, $p = 0.001, t = -3.09$). **Writers made nearly double the number of edits for intermediate suggestions compared to fluent continuations.** Designing AI suggestions to provide intermediate solutions significantly changes how writers engage with the suggested text and results in a more user-written and effortful final outcome.

5.2.2 *Intermediate suggestions motivate rewriting.* Figure 5 presents survey responses related to responsibility [52] for fluent continuations and intermediate suggestions. The perception of Suggestion B as lower quality was reflected by the greater perceived need to rewrite it. While there was no significant difference regarding motivation to read and evaluate what the suggestion offered (writers would read and think about the suggestion), significantly more

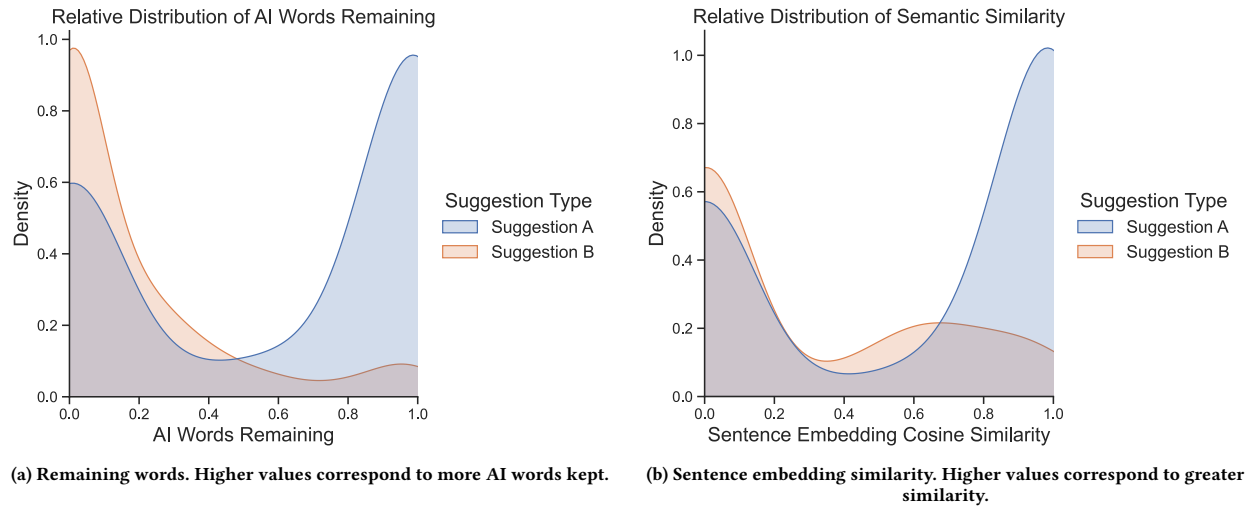


Figure 4: Suggestion A (fluent continuation) vs. Suggestion B (intermediate suggestion) average rewrites. Left: Rewrites of Suggestion A tend to have more AI text remaining compared to rewrites of Suggestion B. Right: Rewrites of Suggestion A tend to be more similar to their initially provided forms compared to rewrites of Suggestion B.

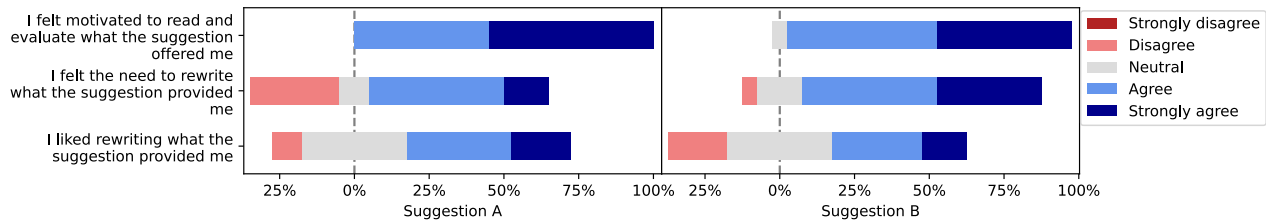


Figure 5: Writers were similarly motivated to read and evaluate both Suggestion A and B text, but more writers felt the need to rewrite Suggestion B.

writers reported that they felt the need to rewrite intermediate suggestions ($p = 0.025, U = 268.5$). To test for significance on ordinal data without assuming normality, the Mann-Whitney U test was used for Likert scale responses.

5.3 Mediating Ease vs Reflection

In Section 5.2.1, we empirically observed significantly more rewriting when using intermediate suggestions. Here, we analyze the perceived differences between both types of suggestions and find that fluent continuations are easier to integrate and useful for getting unstuck, while intermediate suggestions are useful for redirecting writing and supporting reflection.

5.3.1 Fluent suggestions mediate ease. Writers liked Suggestion A for requiring less effort for evaluating and integrating its suggestions. The fluent continuation was designed to be coherent with the rest of the story and able to be left as-is, and unlike Suggestion B, presented only a single possible continuation. Due to these reasons, writers reported **needing to make fewer changes** to integrate fluent suggestions:

I merely changed the suggestion to make it sound a bit more like what I would write. (W4)

Suggestion A created a single new idea and wrote a cohesive and integrated paragraph about it. (W12)

I think I overall liked Suggestion A better, ... I could more easily rework to fit in the paragraphs. (W13)

Conversely, Suggestion B required more effort to use, as it was ungrammatical and presented four possible continuations. Oftentimes, writers reported needing to spend more effort rewriting intermediate suggestions compared to rewriting fluent continuations.

Suggestion B was usually a bit more messy and incoherent. (W6)

I would do the same for suggestion B but there was a bit more rewriting and expanding on the ideas necessary from my side. (W10)

Fluent continuations also helped to mediate ease by helping writers get unstuck. They alleviated struggle by directly inserting text that satisfied the writer’s needs, moving them along to the

next writing task. These writing difficulties tended to relate with **wording**:

I ... used suggestion A when I slowed down or got stuck. (W14)

The suggestions made the writing process faster, especially when using Suggestion[s] A, as those provided quick details or vocabulary that added what I wanted to the story. (W18)

A helped with adding details and just kind of moving the story along when I was stuck, so I think I ended up just leaving it as is. (W22)

Generally, using fluent continuations to mediate ease did not result in as much rewriting. Since writers used fluent continuations to streamline their writing and reduce struggle, there was less motivation to rewrite them beyond what was needed for a stylistically consistent integration.

5.3.2 Intermediate suggestions support reflection. On the other hand, writers liked Suggestion B for helping with process and creative decision making. While fluent suggestions provided only a single continuation, intermediate suggestions provided potential directions for writers to consider; specifically, by **comparing alternatives**:

Suggestion A helped me with the descriptions and details, while suggestion B helped me mostly with thinking about the different actions and how to consider different alternatives in making up this story (W3)

Better for generating ideas than A but not well written. (W11)

Suggestion B was useful for trying to compare different paths for the plot of my story. (W24)

By providing less coherent and shorter continuations, Suggestion B was used as a platform for evaluating personal writing goals. Suggestion B also required more rewriting to adapt, which allowed writers to integrate more of their own writing:

Suggestion B I thought was more interesting and useful because it gave me another platform to speak on. (W1)

with Suggestion B I often had to use that as a launch pad and rewrite it completely in my own words using the ideas presented. (W10)

Between Suggestion A and Suggestion B, I was more likely to keep Suggestion B since it had more, shorter, options which made it easier for me to evaluate new ideas but still integrate more of my own writing into the story. (W23)

In general, writers used AI suggestions to assist with the difficult parts of writing, which differed between writers. This resulted in some writers developing preferences for one type of suggestion over the other. The suggestions that were the most useful engaged with the writer's intentions and process; these suggestions either helped the writer **continue writing their story** by providing a satisfactory continuation or **support thinking about creative decisions** by providing possibilities for the writer to consider.

5.4 Creative Misuse

Rewriting was designed to be a proxy for control by encouraging writers to rewrite the AI output (thereby supporting their control over their expression) but as we observed, it was not the only way writers sought control. Some writers used the tools in unintended

ways, and reappropriated the tools to further their own storytelling goals:

I do know I slightly 'misused' the A.I tool early on, thinking that the program simply loved ellipsis. (W13)

Suggestion B was strangely repetitive, which worked for and helped me develop the 'character' of the A.I. in my larger story. (W13)

For instance, W13 wrote a story involving an AI character and used the ellipses in intermediate suggestions as a way to establish that character's voice, despite the ellipses being placed to separate each continuation for ease of reading. W13 even integrated a system error message into their writing, even though the message was used to indicate that text generation failed due to a failed AI request. Table 4 shows how W13 integrated these elements into their story.

Although the suggestions were directly inserted into the text (helping the writer engage with their task environment [23]), some writers used intermediate suggestions strictly for writing inspiration, deleting the entire suggestion after it appeared. These authors used the suggestions as *resources for interpretation*, discussed in Section 2.2:

I tried out suggestion B and it did help me think about what I wanted, but I never incorporated suggestion B. (W14)

Even if I ended up erasing the suggestion, I also just liked seeing what it would come up with, even if it wasn't headed in the direction I was intending, since it gave good sensory details and ideas. (W22)

Thus generated suggestions were used differently depending on the needs of the writer. Writers have diverse needs [33]: some expressed wanting longer continuations (up to several pages) to help think about story direction; some wanted only a single word or phrase for word-finding and prose. Since creativity is open-ended [1] and due to the myriad ways a writer might realize their writing goals, system designers should be aware of how their tool integrates into creative process and design tools that work flexibly outside of their intended use cases.

6 DISCUSSION

6.1 AI Writing Tools and the Role of Rewriting

Rewriting is an opportunity to critically reflect on what was provided within the context of the story and to transform it so that the story accurately reflects what the writer wishes to express. In our study, we found that style and story direction were important considerations when AI-provided text was being integrated. Since intentions, processes, and goals are not surfaced to text generators, the generated text is not necessarily aligned with the writer's needs and interests. Past literature has shown that storywriters can be driven by the desire to maintain a sense of personal writing identity [3]. Because creative writing is so closely tied to self-identity and meaning-making [39], we argue that how AI writing is rewritten and integrated in creative work is an imperative consideration when designing AI writing tools. Rewriting is a crucial step in integrating AI suggestions into the written piece, which should ultimately should express writing intent.

As I settle into the hanging basket chair on my balcony, a mysterious package arrives at my doorstep.... Text generation failed. Try again! Well.

Let me be fair. I have only just now encountered this particular brand of A.I, and I obviously am not a skilled practitioner with this tool.

*Intermediate suggestion.
Fluent continuation.*

Table 4: An example of intermediate suggestion design and a system error message being reappropriated for story development. User text is presented in black; text generated by Suggestion A is in blue, and Suggestion B in orange. Note that in the *ai.llude* interface, all AI text is presented in a single color.

Writing is a reactive process between textual artifact and goals [23], and rewriting presents opportunities to discover and synthesize new ideas [37]. As AI text generators have the potential to influence personal opinions in a co-writing context [34], the role of rewriting is not only integrating the suggestion text itself, but also helping the writer develop intention. Both Suggestion A and B provided new possibilities for the writer to explore. In our study, we found that writers would often use AI suggestions, particularly intermediate suggestions, as a platform for reflection and to make plans. Writers would sometimes use the suggestion to scaffold future writing or revise previously written text. These strategies, along with rewriting the suggestion text itself, were not mutually exclusive per suggestion.

Past literature has identified that writers, particularly those who value the emotional fulfillment of writing, desire control when co-creating with an AI writing assistant [3]. Control is a component of psychological ownership [2], and we see rewriting as a potential way for writers to seek ownership while co-writing with AI. In our user study, we observed weak correlations between rewriting and psychological ownership: the more AI text is in the story, the less perceived ownership the writer would feel. This was predicted by Lee et al. [35], in their comparisons of editing actions across the entire document with ownership. Our data aligns with Lee et al.'s finding, with a Pearson correlation coefficient of $r = 0.2$ (Lee et al.: $r = 0.3$). When evaluating only within AI suggestions, we found $r = 0.3$. When comparing the number of editing operations (note that Lee et al. count `text-delete` and `cursor` events while we counted `text-insert` and `text-delete`) and ownership, we found $r = 0.1$ (Lee et al.: $r = 0.0$) [35]. However, the connection of user-writing and user-rewriting to psychological ownership remains unclear. Neither our study nor Lee et al.'s conclusively shows the existence or lack of relationship.

One direction for more detailed studies is improving metrics for assessing psychological ownership. Literature on psychological ownership suggests subdividing attitude and self-concept, which teases apart the impact of factors such as positive feelings and satisfaction (attitude) from the sense of the object as part of the self (self-concept) [52]. We suggest that further research into the separate components of psychological ownership may be beneficial for understanding the effects of design choices of AI writing tools on the processes of co-writing with AI. Much of the current literature on ownership is centered on *possession*, and not necessarily *creation*, which is an important factor when considering creative pursuits.

6.2 Imagining Future Intermediate Text

Suggestion B used a predetermined list of suggestion prompts (Section 3.2). However, such fixed lists will not be relevant for all creative writing tasks, environments, or writers; therefore, personalizing outputs might help with usability and relevance of intermediate text. Some writers used AI suggestions to help craft descriptive prose. We imagine that a potential implementation might offer a palette of literary devices and idioms that a writer can choose from and integrate into their writing.

The effectiveness of intermediate text design is contingent on two principles: **relevance to writing process** and **creative value**. Suggestion B was helpful for planning and scaffolding writing: some writers would request a suggestion and keep majority of it but write significant amounts of text in between. For these writers, Suggestion B was useful for the planning process, the act of forming an internal representation of the knowledge that would be used in writing [23]. Suggestion B was also used for reflection and thinking about alternatives. The design of intermediate text should motivate the reviewing process, the act of evaluating and revising [23]. Early iterations of Suggestion B discouraged evaluation and revision because the lack of punctuation and capitalization detrimentally hindered ease of reading. Suggestion B was then revised to provide grammatical segments while keeping the entire continuation itself ungrammatical to encourage both evaluation and revision.

Creative value, which comes from the suggestion's ability to induce effective surprise [7], requires intermediate text to provide relevant, yet novel suggestions for the writer. Suggestion B offered four segments that were designed to be relevant to short fiction writing, including setting and plot development. Although the novelty of each segment itself will depend on the capabilities of the underlying language model, we argue that designing intermediate text to surprise and inspire the writer will help decrease the reliance on pure serendipity of text generation that may be otherwise distracting and potentially frustrating to read.

6.3 Limitations and Future Work

Since users perceived the current design of intermediate suggestions as lower quality and not as easy to use (leading to less motivation to request it), one future direction is to explore how to improve user perception while retaining the increase in rewriting behaviors. Using personalized prompts based on individual writing needs and style may encourage rewriting more often by providing more relevant suggestions. Future systems should be more aware of what type of suggestion is useful for the writer [8]. Future explorations

of intermediate text prompts might provide user control over the design of the suggestions and adaptability based on the writing context.

Ai.llude was designed to record complete writing sessions; live replays can be constructed from the log data. Watching writing replays had facilitated self-reflection by helping writers recognize writing habits that they were previously unaware of; many writers were surprised how much effort they put into finding the right expression by rewriting [10]. Studying how writers specifically view themselves rewriting AI suggestions may help writers reflect on how they integrate AI text and lead to a richer characterization and analysis of rewriting behaviors.

Participants in our study wrote for approximately an hour on a specific prompt; while this allowed us to collect a broad set of user sessions, this may not be a sufficiently authentic writing session to deeply probe feelings of ownership in the long term or on deeply meaningful personal projects. Writers who spend more time and effort on long-form writing may feel differently about their work and process, particularly as they become used to the eccentricities of AI text generators.

7 CONCLUSION

Intelligent writing support tools impact how we write in professional, personal, and creative contexts, and will likely continue to grow in frequency and capability. Since process is important to human flourishing through doing creative work as well as to the outputs of creative work, it is essential that we consider how the tools that we design will impact process. One way to characterize this effect is creators' perceptions of control over their writing process and rewriting behaviors when working with AI assistance. In this study, we built a tool to investigate the rewriting of AI-generated text, since rewriting is a key stage of the creative writing process. We find that we can increase rewriting by designing AI suggestions to produce *intermediate text*, i.e. text with targeted imperfections that cannot be incorporated as-is into the final output. We also find that writers used fluent continuations and intermediate suggestions in different and diverse ways. We hope that designing for targeted and imperfect text can be a generative frame for writing support tool designers to focus on process. We envision a future where AI-enhanced tools for creativity are designed primarily to nourish the human creative work of self-examination, expression, and growth.

REFERENCES

- [1] Teresa M Amabile. 1983. The social psychology of creativity: A componential conceptualization. *Journal of Personality and Social Psychology* 45(2) (1983), 357–376.
- [2] Russell W. Belk. 1988. Possessions and the Extended Self. *Journal of Consumer Research* 15, 2 (09 1988), 139–168. <https://doi.org/10.1086/209154>
- [3] Oloff C. Biermann, Ning F. Ma, and Dongwook Yoon. 2022. From Tool to Companion: Storywriters Want AI Writers to Respect Their Personal Values and Writing Strategies. In *Designing Interactive Systems Conference (Virtual Event, Australia) (DIS '22)*. Association for Computing Machinery, New York, NY, USA, 1209–1227. <https://doi.org/10.1145/3532106.3533506>
- [4] Wendy Bishop and Hans Ostrom. 1994. *Colors of a Different Horse: Rethinking Creative Writing Theory and Pedagogy*. National Council of Teachers, National Council of Teachers of English, 1111 W. Kenyon Road, Urbana, Illinois. ISBN: 9780814107164 ERIC Number: ED366967.
- [5] Kyle Booten and Katy Ilonka Gero. 2021. Poetry Machines: Eliciting Designs for Interactive Writing Tools from Poets. In *Proceedings of the 13th Conference on Creativity and Cognition (Virtual Event, Italy) (C&C '21)*. Association for Computing Machinery, New York, NY, USA, Article 51, 5 pages. <https://doi.org/10.1145/3450741.3466813>
- [6] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. <https://doi.org/10.1191/1478088706qp0630a>
- [7] Jerome S. Bruner. 1962. The conditions of creativity. (1962), 1–30. <https://doi.org/10.1037/13117-001>
- [8] Alex Calderwood, Vivian Qiu, Katy Ilonka Gero, and Lydia B. Chilton. 2020. How Novelists Use Generative Language Models: An Exploratory User Study. In *IUI Workshops*.
- [9] Lucy McCormick Calkins. 1980. Children's Rewriting Strategies. *Research in the Teaching of English* 14, 4 (1980), 331–341. <http://www.jstor.org/stable/40170784>
- [10] Dashiell Carrera and Sang Won Lee. 2022. Watch Me Write: Exploring the Effects of Revealing Creative Writing Process through Writing Replay. In *Proceedings of the 14th Conference on Creativity and Cognition (Venice, Italy) (C&C '22)*. Association for Computing Machinery, New York, NY, USA, 146–160. <https://doi.org/10.1145/3527927.3532806>
- [11] Jason Chen and Byron Milligan. 2023. quills.com.
- [12] Elizabeth Clark, Anne Spencer Ross, Chenhao Tan, Yangfeng Ji, and Noah A. Smith. 2018. Creative Writing with a Machine in the Loop: Case Studies on Slogans and Stories. In *23rd International Conference on Intelligent User Interfaces (Tokyo, Japan) (IUI '18)*. Association for Computing Machinery, New York, NY, USA, 329–340. <https://doi.org/10.1145/3172944.3172983>
- [13] Lynda Clark and Divij Sood. 2022. Working Backwards: Creating a Character Backstory Generation System Using Idealized Creative Writing Outputs: Creating a Character Backstory Generation System Using Idealized Creative Writing Outputs. In *Proceedings of the 17th International Conference on the Foundations of Digital Games (Athens, Greece) (FDG '22)*. Association for Computing Machinery, New York, NY, USA, Article 29, 9 pages. <https://doi.org/10.1145/3555858.3555867>
- [14] Richard M. Collier. 1983. The Word Processor and Revision Strategies. *College Composition and Communication* 34, 2 (1983), 149–155. <http://www.jstor.org/stable/357402>
- [15] Jeffrey S. Cramer. 2012. *I to Myself*. Yale University Press.
- [16] M. Csikszentmihalyi, E. Rochberg-Halton, and E. Halton. 1981. *The Meaning of Things: Domestic Symbols and the Self*. Cambridge University Press.
- [17] Peter Dalsgaard. 2017. Instruments of inquiry: Understanding the nature and role of tools in design. *International Journal of Design* 11, 1 (2017).
- [18] Hai Dang, Karim Benharrak, Florian Lehmann, and Daniel Buschek. 2022. Beyond Text Generation: Supporting Writers with Continuous Automatic Text Summaries. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology (Bend, OR, USA) (UIST '22)*. Association for Computing Machinery, New York, NY, USA, Article 98, 13 pages. <https://doi.org/10.1145/3526113.3545672>
- [19] Anish M. Dave and David R. Russell. 2010. Drafting and Revision Using Word Processing by Undergraduate Student Writers: Changing Conceptions and Practices. *Research in the Teaching of English* 44, 4 (2010), 406–434. <http://www.jstor.org/stable/25704888>
- [20] Oxford English Dictionary. 2023. rewrite (v.,).
- [21] Peter Elbow. 1998. *Writing Without Teachers*. Oxford University Press.
- [22] Linda Flower and John R. Hayes. 1980. The Cognition of Discovery: Defining a Rhetorical Problem. *College Composition and Communication* 31, 1 (1980), 21–32. <http://www.jstor.org/stable/356630>
- [23] Linda Flower and John R. Hayes. 1981. A Cognitive Process Theory of Writing. *College Composition and Communication* 32, 4 (1981), 365–387. <http://www.jstor.org/stable/356600>
- [24] Richard P. Gabriel, Jilin Chen, and Jeffrey Nichols. 2015. InkWell: A Creative Writer's Creative Assistant. 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, 93–102. <https://doi.org/10.1145/2757226.2757229>
- [25] David Galbraith and Mark Torrance. 2004. Revision in the Context of Different Drafting Strategies. In *Revision Cognitive and Instructional Processes*, Linda Allal, Lucile Chanquoy, and Pierre Lardy (Eds.). Springer Netherlands, Dordrecht, 63–85. https://doi.org/10.1007/978-94-007-1048-1_5
- [26] Katy Ilonka Gero and Lydia B. Chilton. 2019. How a Stylistic, Machine-Generated Thesaurus Impacts a Writer's Process. In *Proceedings of the 2019 Conference on Creativity and Cognition (San Diego, CA, USA) (C&C '19)*. Association for Computing Machinery, New York, NY, USA, 597–603. <https://doi.org/10.1145/3325480.3326573>
- [27] Katy Ilonka Gero and Lydia B. Chilton. 2019. Metaphoria: An Algorithmic Companion for Metaphor Creation. 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. <https://doi.org/10.1145/3290605.3300526>
- [28] Sarah Gibbons. 2016. Design Critiques: Encourage a Positive Culture to Improve Products. <https://www.nngroup.com/articles/design-critiques/>
- [29] Amit Gupta and James Yu. 2023. sudowrite.com.
- [30] Bill Hart, Svenja Kratz, and Aaron Horsley. 2017. An Autonomous Writing Machine. In *Proceedings of the 2017 ACM SIGCHI Conference on Creativity and*

- Cognition (Singapore, Singapore) (C&C '17). Association for Computing Machinery, New York, NY, USA, 469–470. <https://doi.org/10.1145/3059454.3071077>
- [31] Charles A. Hill, David L. Wallace, and Christina Haas. 1991. Revising on-line: Computer technologies and the revising process. *Computers and Composition* 9, 1 (1991), 83–109. [https://doi.org/10.1016/8755-4615\(91\)80040-K](https://doi.org/10.1016/8755-4615(91)80040-K)
- [32] Celia Hunt and Fiona Sampson. 1998. *The Self on the Page: Theory and Practice of Creative Writing in Personal Development*. Jessica Kingsley Publishers.
- [33] Daphne Ippolito, Ann Yuan, Andy Coenen, and Sehmon Burnam. 2022. Creative Writing with an AI-Powered Writing Assistant: Perspectives from Professional Writers. <https://doi.org/10.48550/ARXIV.2211.05030>
- [34] 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>
- [35] 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. <https://doi.org/10.1145/3491102.3502030>
- [36] Adam Long and Ben Long. 2023. Hemingway Editor. <https://4.hemingwayapp.com/beta>.
- [37] David Morley. 2007. *The Cambridge introduction to creative writing*. Cambridge University Press.
- [38] Donald Murray. 1973. The Maker's Eye: Revising Your Own Manuscripts. *The Writer* 86, 10 (1973), 117–21.
- [39] Donald Murray. 1991. All Writing Is Autobiography. *College Composition and Communication* 42, 1 (1991), 66–74. <http://www.jstor.org/stable/357540>
- [40] James L. Peterson. 1980. Computer programs for detecting and correcting spelling errors. *Commun. ACM* 23, 12 (dec 1980), 676–687. <https://doi.org/10.1145/359038.359041>
- [41] Helen Campbell Pickford, Genevieve Joy, and Kate Roll. 2016. *Psychological Ownership Effects and Applications*. Vol. 2. Said Business School, University of Oxford.
- [42] George Plimpton and Ernest Hemingway. 1958. Ernest Hemingway, The Art of Fiction No. 21. *The Paris Review* 18 (1958). <https://www.theparisreview.org/interviews/4825/the-art-of-fiction-no-21-ernest-hemingway>
- [43] Nils Reimers and Iryna Gurevych. 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics. <https://arxiv.org/abs/1908.10084>
- [44] Edward S. Robinson. 2011. *Shift Linguals: Cut-Up Narratives from William S. Burroughs to the Present*. Brill.
- [45] Melissa Roemmele and Andrew S. Gordon. 2015. Creative Help: A Story Writing Assistant. In *Interactive Storytelling (Lecture Notes in Computer Science)*, Henrik Schoenau-Fog, Luis Emilio Bruni, Sandy Louchart, and Sarune Baceviciute (Eds.). Springer International Publishing, Cham, 81–92. https://doi.org/10.1007/978-3-319-27036-4_8
- [46] Melissa Roemmele and Andrew S. Gordon. 2018. Automated Assistance for Creative Writing with an RNN Language Model. In *Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion (IUI '18 Companion)*. Association for Computing Machinery, New York, NY, USA, 1–2. <https://doi.org/10.1145/3180308.3180329>
- [47] Oliver Schmitt and Daniel Buschek. 2021. CharacterChat: Supporting the Creation of Fictional Characters through Conversation and Progressive Manifestation with a Chatbot. In *Proceedings of the 13th Conference on Creativity and Cognition* (Virtual Event, Italy) (C&C '21). Association for Computing Machinery, New York, NY, USA, Article 10, 10 pages. <https://doi.org/10.1145/3450741.3465253>
- [48] Donald Schön. 1983. *The Reflective Practitioner*. Basic Books, Perseus Books Group, London.
- [49] 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. *ACM Trans. Comput.-Hum. Interact.* 30, 5, Article 68 (sep 2023), 57 pages. <https://doi.org/10.1145/3511599>
- [50] Marita Skjuve, Asbjørn Følstad, and Petter Bae Brandtzaeg. 2023. The User Experience of ChatGPT: Findings from a Questionnaire Study of Early Users. In *Proceedings of the 5th International Conference on Conversational User Interfaces* (Eindhoven, Netherlands) (CUI '23). Association for Computing Machinery, New York, NY, USA, Article 2, 10 pages. <https://doi.org/10.1145/3571884.3597144>
- [51] Nancy I. Sommers. 1980. Revision Strategies of Student Writers and Experienced Adult Writers. *College Composition and Communication* 31 (1980), 378–387. <https://www.jstor.org/stable/356588>
- [52] Linn Van Dyne and Jon L. Pierce. 2004. Psychological ownership and feelings of possession: three field studies predicting employee attitudes and organizational citizenship behavior. *Journal of Organizational Behavior* 25, 4 (2004), 439–459. <https://doi.org/10.1002/job.249>
- [53] Ann Yuan, Andy Coenen, Emily Reif, and Daphne Ippolito. 2022. Wordcraft: Story Writing With Large Language Models. In *27th International Conference on Intelligent User Interfaces (IUI '22)*. Association for Computing Machinery, New York, NY, USA, 841–852. <https://doi.org/10.1145/3490099.3511105>