# Social Dynamics of AI Support in Creative Writing

Katy Ilonka Gero
katy@cs.columbia.edu
Columbia University
New York City, New York, USA

Tao Long
long@cs.columbia.edu
Columbia University
New York City, New York, USA

Lydia Chilton
chilton@cs.columbia.edu
Columbia University
New York City, New York, USA

## ABSTRACT

Recently, large language models have made huge advances in generating coherent, creative text. While much research focuses on how users can interact with language models, less work considers the social-technical gap that this technology poses. What are the social nuances that underlie receiving support from a generative AI? In this work we ask when and why a creative writer might turn to a computer versus a peer or mentor for support. We interview 20 creative writers about their writing practice and their attitudes towards both human and computer support. We discover three elements that govern a writer's interaction with support actors: 1) what writers desire help with, 2) how writers perceive potential support actors, and 3) the values writers hold. We align our results with existing frameworks of writing cognition and creativity support, uncovering the social dynamics which modulate user responses to generative technologies.

## CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**; **HCI theory, concepts and models**; • **Computing methodologies** → **Natural language processing**.

## KEYWORDS

creative writing; writing support tools; writing assistants; human-AI collaboration; language models

**ACM Reference Format:**
Katy Ilonka Gero, Tao Long, and Lydia Chilton. 2023. Social Dynamics of AI Support in Creative Writing. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3544548.3580782

## 1 INTRODUCTION

A writer has been sitting at her desk for hours, working on her science fiction novel. At some point, her mind starts to wander. Instead of checking her email and getting distracted, she turns to a computer program, which generates the next few sentences of her chapter. She's always curious where the computer will take the scene, and this re-engages her in the writing process.

A different writer is working on a short story about a young girl in Calcutta and she wants some feedback on the character development. Sometimes she sends her stories to her younger brother; they share a similar background—both are Indians living in America—and she often sends him work when she wants an Indian perspective. But instead she sends the story to her friend from high school, who likes to write about girlhood and may be better at understanding that aspect of the story.

In each of these situations, a writer is looking for support. Where they look for support is a complicated matter. What kind of support is available to them? What individual characteristics do these support actors have? Does the writer trust the support actor? In this work we study the social dynamics of when and why creative writers request support, whether that support comes from a peer, mentor, or computer program. Though we study a specific endeavor—creative writing—our work has implications for all kinds of complicated tasks and the role of technology in pursuing them.

Our work is inspired by recent rapid improvement in generative AI technologies, which are demonstrating more general and flexible capabilities. Natural language generation in particular has seen a burst of interest as large language models like GPT-3 [7] and OPT [42] have proven able to follow complicated instructions and generate coherent, creative text. In turn, the HCI research community has leapt to the challenge of understanding how users can interact with these technologies, developing new interaction techniques like prompt chaining [40] and documenting new interaction patterns like integrative leaps [34]. However, less ink has been spilled on the social-technical gap [1] that these new technologies pose. What is the nuance and context that underlies collaboration in a creative task? And how does this nuance play out when collaborating with a generative AI?

To respond to these questions, we consider how large language models might fare as writing support tools given the social dynamics of requesting help with a creative writing project. Creative writing is a keystone of human endeavors, accounting for a multitude of cultural artifacts from poetry to journalism. We ask the following research question:

> *RQ: When and why might a creative writer turn to a computer versus a peer or mentor to provide support?*

We interview 20 creative writers from a variety of writing genres. Our interviewees include 6 creative writers currently using a generative AI creative writing support tool. The interviews focus on what influences their writing practice, first asking about existing kinds of influence such as suggestions from peers, then asking about hypothetical computer programs that could provide human-like support. Our work builds on existing frameworks of writing cognition and creativity support, uncovering new dynamics


*CHI '23, April 23–28, 2023, Hamburg, Germany*

ACM ISBN 978-1-4503-9421-5/23/04…$15.00
https://doi.org/10.1145/3544548.3580782

which modulate user responses to technology. Through a qualitative analysis we discover three elements that govern a writer's interaction with potential support actors:

- what writers *desire* help with,
- how writers *perceive* potential support actors, and
- the *values* writers hold about the writing process.

We align our results with two existing models. First, we find that the types of support desired can be aligned with the updated cognitive process model of writing [17], which includes motivation and goals. Second, we build upon the support relationship types proposed in Chung et al. [8], contributing the organizing principles of a) how an artist perceives a support actor, and b) how an artist's values impacts when and where they turn for support.

Finally, we discuss how our findings reveal when and why writers might turn to computers for support, and outline future work for building rich interactive writing support tools.

## 2 RELATED WORK

### 2.1 Models of Writing

In 1981 Flower and Hayes proposed the cognitive process model of writing [10]. Their model, based on empirical research, describes three main cognitive processes—planning, translating, and reviewing—which are nonlinear and hierarchical. Planning includes generating ideas, setting goals, and organizing thoughts. Translating is the act of 'translating' ideas and thoughts into words on the page. Reviewing includes evaluating and revising what has been written. Although a simplified view of writing may consider these processes as linear, in reality writers move between all of them throughout a writing session, and within a single process may call upon all three as sub-processes. These processes interact both with the writer's long-term memory, as well as the task environment, which consists of the rhetorical problem that spurs the writing as well as the text produced so far. This model has remained influential in both writing theory research [20] as well as human-computer interaction research that studies the use of technology in writing [16]. We use these cognitive processes to ensure we ask participants about a range of potential types of support a computer could provide.

Fifteen years later, Hayes proposed a new framework for understanding cognition in writing [17], one that he hoped would be useful for understanding a wider range of writing activities. In addition to the three cognitive processes of the original model, the new model called out relations between cognitive processes, working memory, long-term memory and motivation/affect. Additionally, the task environment is expanded to contain the social environment, which includes collaborators.

The work we present in this paper concerns when writers desire support, and what actions they might take to produce that support. The original three cognitive processes are important in understanding what kind of help writers desire (which cognitive process do they want help with), but do little to help us understand their motivation or the social environment in which support is provided. In this way the new framework does a better job at illuminating the results of this paper, and we use it in the results section to corroborate and contextualize our findings.

### 2.2 Writing Support Tools

Computational writing support has a long history, from early spell-checkers [31] to recent programs which steer story generation [9]. They may be used to reduce the effort of texting [32], aid in nuanced word choice [14], improve emotional writing [29], or help in professional contexts [18]. Our work here is to understand the gap between the nuanced and contextual goals that writers have while writing, and the kind of support computers are able to provide. While typical writing support research starts with some kind of need-finding exercise, and ends with an evaluation, individual systems and studies struggle to identify themes in the ways these systems, as a group, succeed or fail in addressing writers' needs and concerns. Any one of the papers cited above will lead you to a range of responses from writers, often disagreeing on the utility of the same tool. While researchers do their best to run controlled experiments, variation in writers' attitudes—both towards technology and writing itself—continue to pop up as noise. For these reasons, our interviews start with questions about writers' attitudes towards writing and being influenced in their writing in general, and then move on to ask about a variety of hypothetical tools. In this way we address the need to understand underlying dynamics of writing support, which can pave the way to better system design and move comprehensive evaluation.

### 2.3 Natural Language Technologies

Natural language technologies have seen a radical shift over the past five years as improvements across all benchmarks have come with the development of larger language models [7, 33, 42]. Though these models are not significantly different from their smaller counterparts, increasing the scale of both the model and the training data has led to many improvements; for instance these models are capable of many tasks with little to no additional task-specific training. Recent work on writing support tools tends to make use of these pre-trained language models as the underlying technology [9, 16, 22].

However, large language models retain many of the problems of their smaller counterparts. They lack an explicit model of factuality, and thus are prone to generate false information [19]. They can generate toxic text in response to even benign-looking prompts [11, 30]. They remain brittle, performing well in some situations and poorly in others [16]. They are difficult to 'steer', and prompt engineering has become a thriving sub-field aimed at figuring out how to best get a model to do what the user would like [2, 24]. Finally, these models come with a slew of ethical issues, from environmental to socio-political [4].

In this work we don't take any particular stance on these technologies, other than to acknowledge that they exist and that their flexibility opens up new possibilities for technology which previously seemed far-fetched. Although these technologies are not capable of acting as writing partners at the level of a human, they are able to produce human-like outputs in many scenarios, and this inspires us to consider such a technology that is able to work 'at the level of a peer' when asking participants about hypothetical writing tools.

# 3 METHODOLOGY

## 3.1 Study Procedure

Between March and August 2022 we interviewed 20 creative writers about their writing practice and their attitudes towards hypothetical computational language technologies. The interviews were focused on:

- the interviewee's existing writing practice (e.g. "What is a piece of writing you are very proud of? Why?"),
- their existing modes of influence (e.g. "Are there people who currently influence or in the past have influenced your writing? Who? In what ways?"), and
- their response to hypothetical computational tools that could act 'at the level of a peer' (e.g. "If a computer program could suggest places to revise like a teacher or peer could, would you use it? Why or why not?").

Some interviewees had experience with an existing creative writing tool called SudoWrite[1]. SudoWrite was chosen as a popular creative writing tool that makes use of contemporary language technology; although it is difficult to ascertain popularity of such tools, SudoWrite appeared to be, at the time of the study, the most prominent such tool in an American context. SudoWrite is a piece of commercial software that requires a monthly subscription and is marketed primarily as a story writing tool. It uses a large language model as its underlying technology. SudoWrite provides capabilities such as continuing a story where you left off, describing a scene, rewriting according to some guidelines, brainstorming plot points, and feedback. We provide examples of SudoWrite's capabilities in Appendix A.1. Writers with experience with SudoWrite had a modified version of part 3 for the interview that focused on their use of and response to SudoWrite, instead of hypothetical tools.

We do not exclusively recruit writers with experience with a creative writing tool for several reasons. First, creative writing tools tend to focus on novelists (this was also seen in our recruitment of SudoWrite users) and we wanted to understand a variety of creative writing practices. Second, people who use creative writing tools today represent a small and likely biased slice of writers; for instance, writers who are technologically savvy or writers who are attracted to new tools and techniques. Finally, writers with experience with creative writing tools, while bringing a more concrete perspective to AI abilities, are also biased by their experience with a particular tool and thus may envision future abilities or tools differently than those who are have yet to experience these technologies as they exist today.

The interviews were semi-structured; the interviewer asked follow-up questions or skipped some questions in the guideline when appropriate given the content and context of the interview. Questions were also altered to best probe the writer's genre. The guideline used by the interviewer can be found in Appendix A.2.

Interviews were conducted via video chat, were conducted in English, and lasted about an hour. Participants were compensated $50USD for their time. The study was deemed exempt by the relevant ethical review board.

[1]https://www.sudowrite.com

## 3.2 Participant Recruitment

*3.2.1 Definition of Creative Writer.* We recruited anyone who identifies themself as a creative writer. We believe the definition of an 'expert' or 'amateur' creative writer is difficult in a field that has unclear professional delineations. Many successful writers retain full-time jobs as teachers, editors, or in unrelated professions, as few are able to make a living from their writing alone. One potential way to screen participants is to select only participants published in certain venues (this approach is used in [6]). However, using publication by a major publishing house as a metric for expertise will continue to enforce the marginalization of many writers, as major publishing houses repeatedly fail to diversify their writers, editors, and leaders [37]. Another would be to recruit only those with a formal creative writing education (e.g. a Masters of Fine Arts in Creative Writing). However, the cost of these programs can be preventative for many people, and analysis has found that novels written by people with an MFA are not detectably different from those written by people without one [36]. Given these concerns, we recruit widely and allow participants to identify themselves as creative writers. This resulted in a range of participants, from tenured professors of poetry with three critically acclaimed books to writers with only informal education currently submitting their first novel for publication.

*3.2.2 Sampling Method.* We used purposeful sampling for maximum variation because we wanted to identify information-rich cases with the aim of "capturing and describing central themes that cut across a great deal of variation" [28]. Given the context of creative writing, we wanted to capture insight from a variety of writing genres, as we expected that different genres may have different aesthetic and personal concerns. For example, a fantasy writer may have different concerns about ownership and voice than a memoirist or poet. In addition to this, we recruited participants who use an existing writing tool, SudoWrite, that provides human-like responses to writing. This allowed us to gain insight into the dynamics of those currently using a large language model-based support tool. We continued recruiting until we had adequate variation in participants, and saturation of themes.

*3.2.3 Recruiting Procedure.* Participants were recruited through MFA graduate school distribution lists, professional networks, and personal contacts and writing communities. For recruiting SudoWrite users, we reached out to writers who had written about their experience with SudoWrite online, either via personal blog posts, published interviews, or social media. We then used snowball sampling, asking those we found to introduce us to other SudoWrite users.

*3.2.4 Participant Population.* Due to our recruiting procedure, all our participants wrote, for the most part, in English and tended to come from an American context.[2] Table 1 reports the genre and writing education of all participants. The demographics of the participants were: 8 women, 7 men, 1 non-binary, 4 undisclosed; 7 aged 18-25, 5 aged 26-35, 2 aged 36-45, 2 aged 46-55, and 4 undisclosed. Most SudoWrite users that were recruited were novelists, in line

[2]Although we did not explicitly ask about language, one writer mentioned that his first and more proficient language is German, but he prefers to do creative writing in English.

**Table 1: Background of Participants. W prefix in ID stands for 'writer'; S prefix stands for 'SudoWrite user'.**

| ID | Genre | Education |
|---|---|---|
| W1 | Non-fiction | Some classes |
| W2 | Science journalism | Some workshops |
| W3 | Poetry | MFA in poetry |
| W4 | TV comedy scripts | Some classes |
| W5 | Fiction, novels | MFA in fiction |
| W6 | Poetry | MFA in poetry |
| W7 | Poetry | MFA in poetry |
| W8 | Fiction, poetry | MFA in fiction |
| W9 | Analysis essays | Informal |
| W10 | Poetry | MFA in creative writing |
| W11 | Historical fiction, science fiction | Some classes |
| W12 | Poetry, essays, short stories | Some classes |
| S13 | Paranormal cozy mysteries (novels) | Some classes |
| W14 | Personal essays | Student news rooms |
| W15 | Creative non-fiction, poetry | Mostly informal |
| S16 | Novels (fantasy, magical realism) | Some classes |
| S17 | Historical fiction | Informal |
| S18 | Fantasy stories | Some classes |
| S19 | Fantasy and science fiction novels | Informal |
| S20 | Young adult science fiction novels | Some classes |

with who the tool is marketed toward. Although we did not explicitly collect information about publication history, we did ask about participants' writing history ('How long have you been writing?') and education ('What kind of formal or informal education have you had as a writer?'). Six had a Master's in Fine Arts in a writing discipline. Two were full-time writers; several had writing-adjacent professions (e.g. Professor of English or librarian); most had published short pieces (essays, short stories, individual poems); several were attempting to publish their first full-length book (poetry or fiction). Our participants skewed slightly amateur, in that the majority pursued writing outside of their primary career, but we also feel that all our participants pursued writing seriously, in that they were actively pursuing writing projects and/or publication goals.

## 3.3 Analysis and Coding

All interviews were transcribed using an automatic transcription software that kept the original audio aligned with the transcription. We used a general inductive approach for analyzing qualitative data [38] because we wanted to develop a framework of the underlying structure of writers' support experiences. Following this method, two researchers independently read all of the transcripts and identified relevant segments of text. Each researcher assigned each segment of text an initial potential low-level category. Then, through repeated discussion, the researchers reduced category overlap and created shared low-level categories associated. Finally, these low-level categories were collected into high-level categories. The researchers repeatedly met and iterated to construct these categories; during the later meetings a third researcher was present to give further insight into the data. This analysis was started concurrently with the interviews being conducted, such that participant recruitment continued until a saturation of themes was found.

By the end of the analysis process, all relevant text segments were collected and annotated with one or more low-level categories. Although our goal is not to make claims about the relative importance of different categories in creative writers as a whole, we do report the prevalence of different categories in the interviews to provide insight into their occurrence in our interviews.

The construction of categories from the data was driven by the research objective to understand what impacts creative writers' desire to interact with and be influenced by computational writing tools. After the categories were consolidated, the researchers considered how the categories relate to models of writing [10, 17], and theoretical work on creativity support [8]. This consideration did not explicitly influence the creation of the categories, but rather constituted further analysis into the relation between categories, and the meaning and implication of the results.[3]

# 4 RESULTS

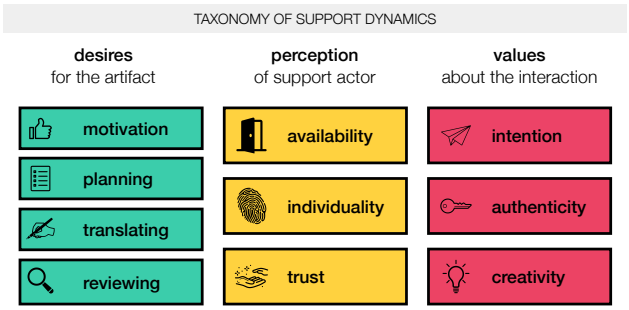


**Figure 1: Results of qualitative analysis.**

Figure 1 shows our three high-level categories: writer desires for support, writer perception of a support actor, and writer values about the writing process. Each category contains 3 or 4 subcategories that delineate what impacts the high-level categories. Our results are not intended to define creative writers' aggregate attitudes toward help seeking or computational tools, but rather account for what impacts their attitudes. Table 2 gives a definition for each category, and exemplary quotes.

## 4.1 Model of Social Dynamics

To better understand the categories in Figure 1, we develop a model that places these categories in relation to each other. This model is an extension of the model presented by Chung et al. [8] that defines types of support relationships for creative practitioners. In Chung et al.'s model, the authors consider three entities: the **artist**, the **actor** (who provides support), and the **artifact** (whatever the artist is trying to create, in our case a piece of writing). First, we adopt their terminology, but replace artist with **writer**, to make clear our participant population. Second, and more importantly, Chung et al. focus their analysis on ways in which artists and actors

[3]Notably the categories for "desires" align with parts of the updated cognitive process model of writing [17]. While the researchers were not attempting to shoehorn the results into this model—for instance, there are many aspects of this model that did not result in categories—knowledge of the model influenced the naming of the categories.

**Table 2: Code Description and Example Quotes.**

| Code | Description | Example Quotes |
|---|---|---|
| **Writer Desires for Support** | | |
| Planning | Coming up with ideas, plotting, deciding on what to work on next. | "I always struggle with plot and endings. I have even tried to use a plot generator. They've been actually not that great." |
| Translation | Figuring out how to 'translate' ideas and thoughts to words on the page. | "My career is based on speed. The faster I write, the faster I can get it out, the more money I'm going to make." |
| Reviewing | Getting feedback, making edits or identifying parts that need work. | "I want other people's perspectives because mine is limited, and people can see things in my poems that I don't necessarily see." |
| Motivation | Getting affirmation, keeping up motivation (on a specific project or in general). | "Often it's at the end of a writing day and I would like to acknowledge that I was productive and I'm excited about something I'd like to share." |
| **Writer Perception of Support Actor** | | |
| Availability | The availability of an actor, for instance an actor must be both physically available (e.g. not asleep), and socially available (e.g writer may feel they've already asked for help too many times). | "Giving feedback takes time and work and I don't want to be asking people to do that work for free all the time." "The advantage that SudoWrite has is availability, because you can't just walk up to a person at any time of the day or middle of the night and go, Hey, I have this idea. How about this?" |
| Individuality | The actor has individual characteristics, such as aesthetic preferences or lived experiences, that modulates the kind of support they provide, and how the writer views any suggestions. | "Every commenter has a perspective, and you understand what they bring to the table. You'd have to develop that about the machine." |
| Trust | The actor must be trusted, for instance to have relevant expertise or to deal with sensitive or personal topics. | "I respect her skills as a writer, and I trust that she knows me well enough to know what's trying to happen; … she's a person who I've given full trust to in many ways and I'm willing to give that trust here." |
| **Writer Values** | | |
| Intention | Writers have intentions or goals that an actor may or may not respect or even understand. | "It's not like the computer can understand what you want to say, they can only see what you have written." "But obviously, you have a reason why to write a story. And I think that is something SudoWrite can't reproduce." |
| Authenticity | Writers have different ideas about what is required to maintain authenticity, and this modulates when and how they want an actor to influence their writing. | "I would be fearful that I wouldn't sound authentic. It's the same reason that I don't really believe in ghostwriting." "I feel like it would feel maybe a little bit creatively dishonest if a computer wrote the ending to my poem for me." |
| Creativity | Writers have different ideas about where creativity lies, and this influences when and how they want an actor to provide support. | "Computers can [only] help us understand or generalize what is the well-trodden path." "[SudoWrite] had introduced a completely new character, and gave him two stanzas of a song that he was singing, it completely wrote the song. And I was floored." |

interact (e.g. by sharing ideas) and ways in which artists and actors contribute to the artifact (e.g. the artist directs the actor to create an artifact). In contrast, our work considers the social dynamics of why and when the artist turns to an actor for support, for instance instead of focusing on the ways an actor might contribute to the artifact, we focus on why an artist would or would not want the actor to contribute in that way. We see our work as an extension to the model of Chung et al.; their model focuses on *types* of support relationships, as defined by kinds of interactions, whereas ours focuses on the *dynamics* of support relationships, as defined by how writers think about these interactions.

In Figure 2 we show a diagram of how the three entities relate to each other given our results. On the left, we see the external dynamics of support. A writer creates an artifact; a writer requests help from an actor; the actor provides support for the creation process. (These actions are not necessarily linear.) Highlighted in red is the request/support dynamic, which is what we are investigating in this study.

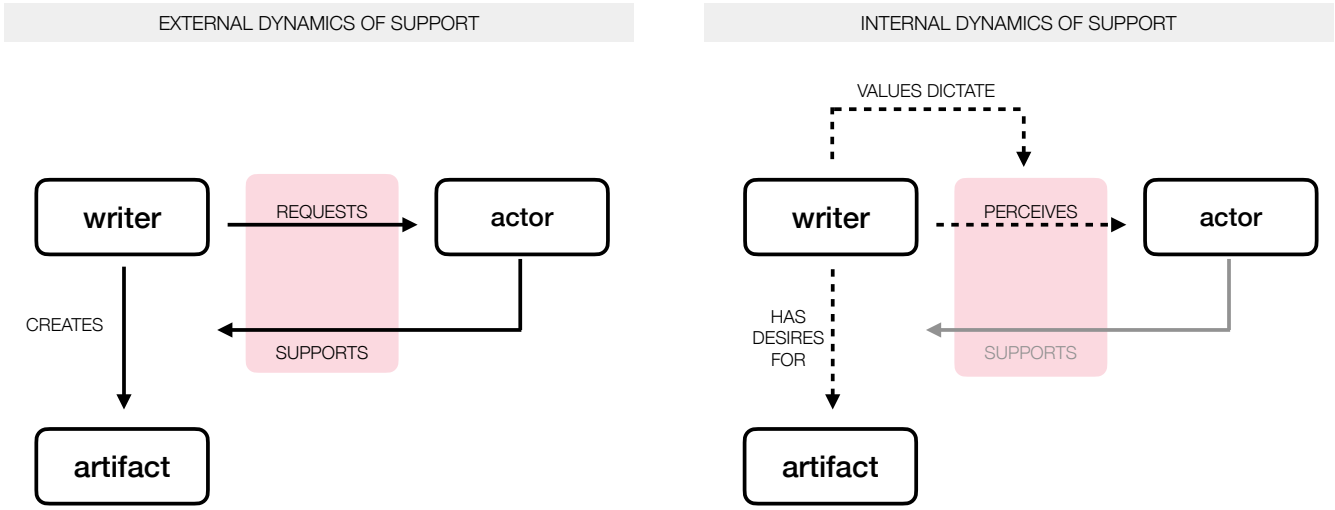


**Figure 2: External and internal dynamics of support.**

On the right side of Figure 2, we see the internal dynamics of support, which we report on in this study. The writer doesn't just create the artifact; they have desires for the artifact, and the actor supports the writer in achieving those desires. The writer also perceives the support actor, and this perception modulates how they choose between different kinds of support. The writer has values about writing more generally, and these values dictate what they want out of the support relationship. In this way, the writer's values impact the kind of support they seek out.

This model is intended to make concrete what goes on 'under the hood' when a writer is looking for help. We call these the social dynamics of support because they model the interactions between individuals—namely, the writer and support actor—and these interactions are the source of aggregate-level behavior, such as a support tool being adopted by many writers.

In the following sections, we report in depth on the themes that emerged from our interviews, and how they together lead to a better understanding of when and why writers look for support.

## 4.2 Writer Desires for Support

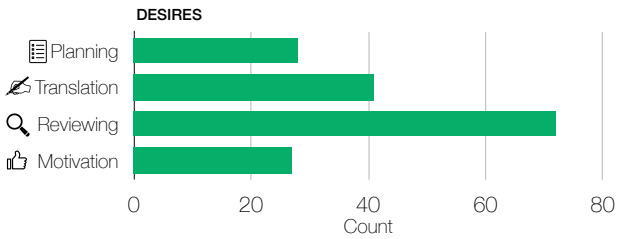


**Figure 3: Prevalence of 'desires' codes in data.**

*4.2.1 Planning.* Writers often sought support for planning. The kinds of planning spanned from early-stage brainstorming before anything had been written, to how to tie up a poem or some plot points. Some talked about wanting open-ended inspiration, while others wanted help with research or coming up with details. S20, working on a science fiction novel, described having a great brainstorming relationship with her teenage son, who also loves science fiction and enjoyed swapping plot ideas.

Many writers talked about doing research as part of their writing. W10 described researching starfish anatomy; W11 talked about researching social mores in the medieval period; W5 talked about referencing a list of slang from a certain historical period. While some writers were hesitant to use a computer for fact-checking, others were open to a computer program that was "kind of like having a research assistant" (W14).

In the interviews we explicitly asked about a hypothetical program that could help with endings. We wanted to understand how writers felt about computational support not just at an early stage, which much research has focused on, but later on in the process. Writers were mixed. W11 said they weren't against such support, but that it would take "the pleasure out of writing. Pleasure is finding the solution to the problem." In contrast, W8 said suggested endings would be really useful because either she would use a suggested ending, or realize that those endings wouldn't work.

*4.2.2 Translation.* The term 'translation' is used in cognitive psychology research on writing to mean the actual act of getting words onto the page (in comparison 'writing' refers to the various cognitive activities involved in writing, many of which don't involve pen to paper or, more likely, fingers to keys). The process of getting words onto the page can be a difficult part of the writing process. Whether it's to start an essay, continue a scene, or just select the correct word for what they wanted to express, all writers discussed the difficulty of translation. However, no writers talked about turning to other people for support in this part of the process. Instead, writers curated their own techniques to get them to write, whether it was to literally "draft standing up" (W2 drafted at a standing desk),

turn to thesauruses (W8 said "sometimes I'm really frustrated because I feel like the words I'm using don't have the right texture or sound"), or turn to a computer program.

Several writers, both those who used SudoWrite and those who didn't, discussed using a computer program to help them with writer's block. W14 described a hypothetical situation, "where I'm not sure how to proceed, but [could use the computer] to see some possibilities laid out." S16 talked about how "it's a lot easier to react to something and make modifications than to come up with something from nothing." S20 talked about how SudoWrite brought ease to her writing, saying she could "pound her head against the wall" and "open up other books and flip through for inspiration", but SudoWrite eliminates this struggle for her when she feels stuck.

*4.2.3 Reviewing.* When talking about how they typically sought out support, writers mostly talked about getting feedback on their writing.[4] Whether they got feedback from a best friend or a professional editor (or, for the lucky few, someone who was both), the ways they talked about feedback were extremely similar, despite our participants coming from a range of genres and writing education. Almost all discussed the importance of needing other people's perspectives, whether on a poem (W10 said "people can see things in my poems that I don't necessarily see") or a newspaper article (W2 said "I was mostly interested in how it was working structurally, and the clarity of the analysis").

Writers expressed the importance of specificity in the feedback. W15 describes how if a reader says, "this might be a bit boring," he may agree but not have enough information to know how to not making it boring, whereas if they said, "this [particular sentence] is where I lost the thread", that would help more with the editing process. W5 described this occurring when his friend gave him some feedback on part of his novel. He describes the situation: "She let me know that it felt contrived to her. And we were able to even pinpoint what language felt contrived to her so that I could then rework it and smooth it into a way that felt more organic." In this situation, W5 had an extensive back and forth with his friend, which resulted in much more helpful feedback.

In this way, writers expected an explanation for feedback, especially if it came from a computer. W5 said, "I would want to be able to interface with the program to understand why it thinks it needs revision." He gave this example of the level of specificity he imagined being incredibly useful: "If it could say, in the whole body of your text, you've only used natural images in your similes. This is an unnatural image that you're drawing from—do you want to have one that breaks that pattern?" W7 similarly wanted to know what the computer was reading for, saying "Maybe if the computer was delineated as reading for technical or imagistic… if I knew what lens it was reading in. Otherwise I'd be like, 'based off of what?'"

Feedback is often laden with implicit value judgments. W1 discussed the situation of a computer giving her feedback that a certain passage is boring, saying "I would be so angry!" because, for example, sometimes the point of a book may be to explore plotlessness. Specificity or being able to question feedback is a way to move past value judgements into more concrete territory. We come back to this idea of value judgements when discussing writers' intentions.

*4.2.4 Motivation.* When writers talked about how they were influenced by other people, or the kind of support they tended to seek out, they often talked about motivation or affirmation as an important part of the writing process. The writers talked about how writing, especially writing a large project like a novel, was a vulnerable, ambitious, and often lonely activity. W4, who was working on the script for a TV show pilot, said, "If I don't periodically get validation? The operation is a bust."

The importance of affirmation ran across most of the writers we interviewed. Writers talked about having no idea how their writing would come across, or wanting something "to confirm it's not trash". S20 described the first time someone else talked about the characters in her novel as if they were real people, and the importance of this indication that what she had created (i.e. the fictional world) was legible to others.

Writers described this as being distinct from feedback used for reviewing or revision. Writers described that when they're looking for support, it can be useful to know or be explicit about if they are looking for help with revision or motivation. As W7 explained, "A lot of times you want validation, but you asked for feedback, or you want feedback, but you're asking for validation."

## 4.3 Writer Perception of Support Actor

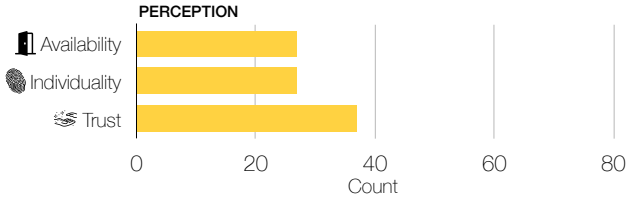


**Figure 4: Prevalence of 'perception' codes in data.**

When writers talked about people who influence their writing, they talked about the particulars of these people. Some talked about their spouses; others talked about family members (parents, siblings, children). There were old high school friends, writing workshop peers, and online writing communities that shared niche interests. When discussing these people, they discussed not just their social relationship, but their perception of how a person might provide support. This perception was key for understanding when and why they might ask a person or a computer for support.

*4.3.1 Availability.* The availability of a support actor was always on writers' minds. Many writers noted that requesting help with their writing was an imposition. By asking for support, they are asking a friend or acquaintance for time, and were always mindful of such a request. W12 preferred reciprocal relationships, where she would also be helping the other person, saying, "If it's not this back and forth, like we're both interested in writing and both giving feedback, then I don't really feel comfortable asking for that favor." Two writers had partners who were heavily involved in their writing. W2 noted, "I'm fortunate to have a partner who's available

[4]Two writers, W3 and W6, rarely sought feedback from peers. Instead, W3 saw successful publication as a kind of useful feedback. W6 talked about how getting lots of feedback was one way to write, but it wasn't the way he was writing.

to offer feedback in a reasonable amount when I'd like it. So I have an in-house editor... Without that, I'd feel much more adrift." But even W5, who talked about his fiancé as an important figure in this writing process, noted that her time was limited: "If she had the time, I would ask her to edit the whole book, but she has a full time job."

Several writers noted that a computer program wouldn't have these issues. S19 said, "The advantage that SudoWrite has is availability, because you can't just walk up to a person at any time of the day or middle of the night and go, Hey, I have this idea. How about this?" Even non-SudoWrite users noted that making use of a computer program wouldn't be an imposition on the computer. W14 said, "I'm imagining things [to request from a computer] that feel too mundane somehow to ask someone for their time."

While computers were understood as being always available, writers didn't necessarily want replicas of their not-always-available peers. W2 said that because he does have access to peers, a program that replicated his peers' response would thus be uninteresting. Instead, he would want a computer program to give him something different, something that wasn't currently available to him.

*4.3.2 Individuality.* When writers discussed a support actor, the details of the actor were important. Not all people were the same, and the individual characteristics of a person (or computer program) impacted not only who they turned to for support, but what they did with the support provided.

One important feature they considered was expertise. W2 talked about how a scientist may be relied upon to point out a factual error, but may not be trusted to critique the quality of an opening paragraph. W9, on imagining a computer reading her work, said, "I would also need to know the reading background. Is this a high school, college, PhD student? What is their level of experience of the topic at hand, so on and so forth? Are they a skeptic or optimist? There's a lot of things to consider."

Writers would come to learn specific characteristics of people that would modulate when they'd turn to these people for support. For instance, W8 discussed how the life experience or writing interests of a peer would dictate who she would send it to, saying "Since my brother is also Indian, if I want to know how something reads to another Indian person, I will show him. But then if I'm writing a story about girlhood, I'll send it to my friend Jen, who also writes about girlhood." S19 even described situations where negative feedback may indicate he's on the right track, saying "I can kind of place their feedback into preference categories… there's certain aspects where I know if a certain person doesn't like that, then it's exactly what I want to achieve in that part of the story."

When discussing a computer support actor, several writers talked about the impossibility of a "universal" reader. W10 worried that computers would represent only a dominant perspective, saying, "The 'universal' perspective has been the perspective of cis straight white men and any other perspective is just not considered universal." Others noted that, based on their understanding of how such a computer program might work, it would reflect generalizations of its training data, and lose the individuality that people provide. W6 described the uniqueness of humans in this way: "Let's just say there's a 1%, that is unpredictable, a response they'll have that does not fit the pattern … I'm interested in that 1%, too. I like the inherent unpredictability of a person."

SudoWrite users were able to articulate the unique characteristics of SudoWrite. S20 describes their sense of SudoWrite:

> A peer is someone who is grounded in a very specific point of view, and culture and identity and preference, you know, their own reading habits and a peer can be a very valuable partner … when I turn to SudoWrite, I know that I'm getting feedback and interactions with my work that is not personal at all … The amount of information on hand that SudoWrite is pulling from is this vast trove. And that's something that a human could never, even if they're well read, could really never achieve.

In this way SudoWrite presented a fundamentally different kind of individuality than a person. SudoWrite is more general, but also more well-read and capable. Still, SudoWrite users discussed the strengths and weaknesses of SudoWrite as they might for a person. For instance, S13 talked about SudoWrite's excellent ability to write descriptors into a scene, but noted that it's terrible at humor.

*4.3.3 Trust.* When discussing when and why writers wanted support, trust repeatedly came up as an important modulating factor. The idea of a 'trusting' relationship with a peer or mentor was key in several respects, and represented a range of interconnecting themes about how the writer perceived a support actor.

Sometimes trust had to do with privacy. W7 talked about how sharing her work with someone meant she's "going to show them a very deep inner part of [her] mind", and mentioned that there may be times when you don't want to share something "extremely sensitive" with a person yet. In this situation, a computer may be a useful intermediary step as a kind of anonymous support actor.

Relatedly, trust also had to do with vulnerability. Many writers talked about the emotional difficulty of getting feedback on their work, and that sometimes they were looking for affirmation, while other times they wanted a more critical response. W10 says, "If it's a poem that is really close to my heart, sometimes if I get feedback on it, I don't even look at the feedback for a couple of days until I'm ready to receive it." W14 talked about the fear of not being good enough, and noted that a computer seemed to not trigger this kind of vulnerability. Her comments about this articulate in general the kind of fears writers have when getting support:

> I think I probably wouldn't feel self conscious [with a computer]. On the basis that I wouldn't have an ongoing social relationship outside of the editing relationship. I think the basis of a lot of self consciousness is maybe the hope, or the anticipation, that we might have some kind of interpersonal relationship, even in passing as acquaintances.

At other times trust was about how well the support actor understood the *writer's* unique characteristics. Someone who had read their work before, who understood their writing 'crutches' and where they tend to be strong, would give better feedback than someone who didn't know anything about them. W15 talked about how he gave more weight to someone who had seen a lot of his

work before, even if he didn't like their feedback or was skeptical for some reason.

Finally, trust often had to do with respect and admiration. W10 talked about sending her manuscript "to the editor-in-chief of that press, because he has worked on a ton of books." W15 talked about a "deference to expertise." When discussing trust, writers often referred to many of these different aspects coming together to create a trusting relationship.

## 4.4 Writer Values

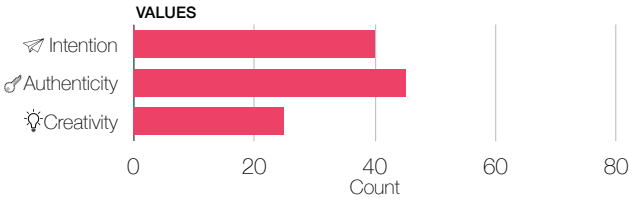


**Figure 5: Prevalence of 'values' codes in data.**

The writers varied on what they value when it comes to writing. For instance, some writers thought never having writer's block was key to their identity as a writer, while others were content to experience this common plague. The idea that some parts of writing might be more important than others, and that writers disagree on what those parts are, is key to understanding when and why a writer turns to support, and what kind of support they'd like to have. There were three elements that writers continually brought up in the interviews, which we discuss here.

*4.4.1 Intention.* The best support came from those who understood the writer's intention. Writers talked about writing as being inherently subjective, and so their goals guided how a writer would (or would not) incorporate feedback or suggestions from peers and mentors. W4 described this as deciding "whether the feedback resonates with me or not". Others noted that suggestions were just suggestions, and they always had the authority to accept or reject them. Part of this was due to the subjective nature of writing; there is no 'ground truth' for good writing, and different writers are often trying to do different things. S19 used the metaphor of designing a roller coaster:

> Plato might not be as much fun as a rollick in a French novel, but they're all ultimately about asking a reader to engage and go on the ride with the creator. And what I think is super important and interesting is that there has been such meticulous decision making about how that ride is going to be shaped.

Due to this subjectivity, W1 talked about the importance of sometimes ignoring other people's opinions: "there is always a part where I stop and step back and ask if this is what I personally want to say. Divorced from 'is this what people want to hear'."

Support actors wouldn't always understand their intention. Writers talked about the difficulty of finding good readers who understood or respected their personal vision for a piece. Sometimes they'd have to discuss with a reader what they were trying to do in the writing, in order for the reader to know if they achieved that.

When it came to computers, this idea of sharing the intention was often a roadblock. W1 talked about how "it's not like the computer can understand what you want to say; they can only see what you have written." This is related to the idea of individuality in the support actor—some actors understand you better, or have a shared context—but relates specifically to the understanding of your intention, which may be unique, personal, and hard to describe. S19 noted that a computer does not bring the same kind of intention that another person does. "Obviously, you have a reason why you are writing a story. And I think that is something SudoWrite can't can reproduce." W12 saw this as a benefit of working with a computer, noting that it's easier to "stay on track with a creative vision" if the computer doesn't have a creative vision of its own.

Intention is closely related to Chung et al.'s support relationship types [8], and theories of co-creativity more generally. Not a single writer in our study considered a human or computer support actor, even hypothetically, to be a leader (or even co-leader) in the writing process. Instead, support actors were always subordinate to the writer's intention.

*4.4.2 Authenticity.* Writers talked about authenticity, or their 'voice', as a concern when it came to incorporating the ideas or suggestions of others. Here, we describe four types of authenticity issues that came up in our interviews: 1) the reader's sense of authenticity, 2) the impact of even viewing suggestions, 3) differing opinions on where authenticity lies, and 4) human v. computer authenticity issues. These four themes shed new light on the problem by addressing more specific concerns than authenticity generally.

First, some writers worried not about their own sense of authenticity, but the sense of their reader. Would the reader notice that the writing was not always in the writer's own voice? W11 said that readers were very perceptive, and that using a computational helper may be similar to writing in an unfamiliar genre: "if you tried to strain to a genre that you may be not experienced in or you're not that interested in, the audience might be able to suspect that." This, he said, was the same reason he also didn't believe in using ghost writers—he said the reader can tell when the author is not being authentic.[5]

Second, writers worried not just about ownership of words on the page, but how even viewing suggestions might derail them. W1 commented that, "once something is on the page it becomes just a bit harder to imagine what else it can be." W12 noted that the longer she looks at something, the more she likes it, worrying that in moments of difficulty she would come to unconsciously or unintentionally prefer the computer suggestion.

Third, writers had different ideas about which parts of the writing process were most central to their feeling of authenticity. W10 talked about how "the ending is such a key piece of a poem that to have to have a computer do it would feel like cheating", while others were eager to get help with endings. S20 considered coming up with the storyline to be "a very human process", but was happy to use computers for overcoming writer's block. Conversely, W5 said overcoming writer's block and writing every day "makes me

[5]Unknowingly responding to this, S13, a professional genre fiction writer, noted that her beta readers never noticed when she started using SudoWrite. Some beta readers would even comment on phrases they particularly liked, and these phrases would be phrases written by SudoWrite. The writer was shocked.

feel more like a writer" and would never ask anyone or thing to help him with that.

Fourth, some writers had existing close relationships with other writers, and these writers often left strong marks on their writing. In describing what was different between her brother influencing her writing and a computer, W8 said, "it somehow feels like there's more of me in it, maybe because we have a relationship with each other." W5 discussed how he gives trust to his fiancé in a multitude of ways, and he can't imagine giving that same trust to a computer.

*4.4.3 Creativity.* The question of if and how computers can be creative has been explored by researchers [5, 25] and critics [26] alike. We found writers bringing up this question when considering when they would feel comfortable with a computer influencing their writing. Connected to the idea of authenticity, writers considered if a computer could be creative, and how creative computational influence might impact the authenticity of their work.

Writers disagreed on the subject. W1 proposed that what computers produce comes out of what they have seen (i.e. training data) and thus computers "help us understand or generalize what is a well-trodden path". W11 thought a computer would be better at historical novels than science fiction ones for this same reason; he trusted the computer to understand the past but not predict the future. W6 said, "I sort of hold the line in believing that there is something irreducibly human" that a computer could never replicate. Similarly W10 said, "I think very highly of all my friends' creative brains. But I don't think that way about a computer."

Others, especially SudoWrite users, were more positive about the creative potential of a computer. W10 even said she may feel bad compared to the computer's abilities, saying "I feel like it would bruise my ego if I couldn't figure out how to end one of my poems, but a computer came up with this really great idea." S19 related SudoWrite to role playing games that used dice rolls to trigger an idea for world building, saying SudoWrite is essentially a more evocative version of that. S20 discussed a moment when she had SudoWrite complete a scene she had started writing, and it "introduced a completely new character, and gave him two stanzas of a song that he was singing... I was floored." Even still, S20 said her creativity is her "humanity" and was not worried about a computer replacing that. She thought writers should be open to new technology, and make use of whatever was available to them.

W4, a TV comedy writer, had a unique perspective on the creative potential of computers. She discussed the difficulty of writing humorous scenes when people have seen so much TV already. She imagined using a computer to push herself and her writing further, saying, "I might take a first stab at a scene, then I would give the computer the first bit and be like 'you write it'. And then if I wrote the same thing as the computer, I'd know I have to do better." She explained further, "I feel like comedy rests on surprise. So I wouldn't trust a computer to do that. Or if the computer did do that, then there has to be something better."

Overall, many writers thought computers would struggle to be creative, and most held the perspective that, even if computers were to achieve the creativity they imagined, there would always be something irreducibly human that would distinguish their work from that of a computer.

# 5 DISCUSSION

The social-technical gap describes the space between human-human dynamics (highly flexible, nuanced, and contextualized) and human-computer dynamics (rigid, brittle, and unchanging) [1]. In this work, we sought to answer the question *When and why might a creative writer turn to a computer versus a peer or mentor to provide support?* and outlined the rich and sophisticated social dynamics between writers and support actors. Here we synthesize some key ideas. We first outline how the themes of **trust**, **individuality**, and **authenticity / creativity** impact writers' relationship to AI writing tools. Then, we highlight design guidelines, drawing on the themes of **individuality**, **intention**, and **authenticity** to suggest ways that systems can better align with writers' wants and values.

## 5.1 How Writers Relate to AI Writing Tools

*5.1.1 The nuance of trust: writers don't feel self-conscious with computers, letting computers actually get closer to the writer's process.* Writers worried about the relationships they with the people who gave them help. Was the writer asking for help too often? Was the writer's question too obvious, or their concern too insecure? Would the person helping them think the writing was bad? Comparatively, writers were not concerned about maintaining a good relationship with a computer program, and weren't worried about judged. Computers may best serve writers in tasks that feel too mundane, too frequent, or too worrisome to turn to a peer or mentor. This needn't just be spell-check; many writers were interested in programs that could suggest potential endings, point out possible problems, or provide new ideas. But likely computers will be considered a kind of pre-cursor to human support, perhaps even providing a very private exchange. As W14 said, "[a] computer program almost feels like me being in conversation with myself".

*5.1.2 Everything has individuality, but the personal relationships writers have with peers will be difficult, if not impossible, to mimic with computers.* Writers developed relationships with those who gave them help, whether it be the long-term relationship of a sibling, the blooming relationship of a new friend, or the structured relationship of a hired editor. While SudoWrite users did develop an *understanding* of the computational support, they expressly stated it was not personal. Computers will likely struggle to provide the individuality that personal relationships entail, like that of W10's brother who was also an Indian living in America.

In the context of language models, this idea has nuance—many researchers think of pre-trained language models as general purpose[6] while in fact they do represent a particular and unique viewpoint based on their training data. Since most training data is scraped from the web, 'general purpose' models typically reflect white, western, and male perspectives, as these are the highest contributors to textual content on the web. Making explicit these assumptions, as well as highlighting when a model is trained to produce a different perspective (e.g. a model trained on woman-written science fiction, or contemporary Latin American poetry), will be an important part of users developing strong mental models of a system. But it won't necessarily result in writers seeing these models as trusted readers. Computational models, according to our interviewees, are

[6] GPT-3 [7] stands for General Purpose Transformer, v3.

inherently reductive. A writer may get help from a friend who is queer, and this friend may be trusted to read a story from a queer perspective. But that friend doesn't represent all queer people, and writers likely are not interested in computers claiming to represent certain identity groups. It is in this way that computers will struggle to provide writers with personal relationships.

*5.1.3 Writers' attitudes towards authenticity and creativity will impact the kind of computational help they seek out.* Writers had a wide variety of ideas about which parts of writing they considered most central to their creative practice. While some valued the challenge and problem solving of figuring out an ending, others were eager to get help with a concluding paragraph. While some thought writing a thousand words every morning was key to their identity as a writer, others didn't see the point in struggling if computational help was available. This variation seemed to be a function not just of authenticity (what was important that the writer themselves do) but also of creativity (what was central in the creative process). This kind of variation reminds us that writers are not a monolith, and there is no one set of values all writers ascribe to.

We see this variation as an explanation, or perhaps a confounding factor, in previous studies of AI writing tools (e.g. [15, 16, 22]) where the responses to tools can vary widely, and is often unexplained by system quality. While the quality of suggests certainly impacts users' perceptions of a system, they alone will not explain why writers can respond so differently to the same suggestions. Variation in a user base may seem like an obvious conclusion, but we rarely acknowledge this kind of variation when conducting user studies. The variation we saw did not cleanly split across common recruitment demographics like amateur or expert, and should be considered when designing and analyzing studies with writers.

## 5.2 Design Guidelines

*5.2.1 Contextualize the individuality of AI writing system.* Writers are coming to AI systems with a variety of preconceived notions. System designers should consider directing writers' assumptions about the system by providing context, such that writers are more likely to understand system capabilities correctly. This context could be in the form of suggesting a mental model to the writer. A number of caricatures or models of AI writing have been proposed, like thinking of the AI as a 'deranged but very well-read parrot' [35] or considering computer-generated text as a kind of automatic writing [27]. Research on chatbots showed that modulating the metaphor of the bot (e.g. telling the user it was modeled after a toddler versus a trained professional) greatly changed how participants viewed the chatbot's capabilities, even though all participants talked to the same chatbot [21]. But context could also come from giving examples of what a system is and isn't good at or interface design decisions (e.g. where computer-generated text is situated in relation to where the writing occurs).

*5.2.2 Consider writers' intentions.* System designers are obviously trying to build systems that writers find useful. That said, systems typically focus on particular tasks, like describing a setting or sketching a storyline [9]. While this is a useful guiding principle, it can be easy to lose track of the reason a writer is writing in the first place. Intention—what a writing is trying to achieve in their work or why they are pursuing a project—dictates how writers make choices about incorporating suggestions or feedback. As a simple example, if someone critiques a sentence as being cliché, but the writer intended to employ cliché, the critique isn't helpful. Intention may be incorporated into system design in a variety of ways. At a high level, systems could explicitly ask writers to articulate their goals, and attempt to use this articulation to direct writers to particular features or customize system outputs. At a more low level, systems could add 'intention' context to outputs. For instance if a system provides three sentence completion suggestions, and two are visually descriptive and the third includes character action, these suggestions could be colored according to their intent (namely, adding visual description and furthering the action).

*5.2.3 Designing for authenticity.* Writers were concerned with producing work they felt was authentic, or represented their own voice. Authenticity is often related ownership, and research has found that ownership over a final output decreases with the amount of text the system, versus the writer, produced [22]. However, our interviews suggested a more nuanced view of authenticity that is not necessary directly correlated with ownership or levels of contribution. For instance, writers worried about reader perception of authenticity, suggesting that a system that can accurately mimic a writer's unique style may not create authenticity concerns. Similarly, writers discussed how having an extensive relationship with someone meant that that person's contributions felt more authentic; systems that customize to a writer over time may also result in contributions that feel more like they came from the writer (as the writer contributed, implicitly or explicitly, to customizing the system) versus a generic program.

# 6 FUTURE WORK

In this section, we outline ideas for designing new AI writing support tools, as well as important questions to be asking and studying when it comes to human-AI collaboration. In particular we a focus on system ideas that are particularly novel or under-explored in commercial and research systems, and consider research questions we should be tackling to better understand how writers are and will be working with AI systems.

## 6.1 Novel Kinds of AI Writing Support

*6.1.1 Feedback with specificity.* When writers talked about how other people influenced their writing, they almost always talked about feedback; no participant discussed other people writing for them. Despite this, most writing support tools studied in HCI generate text for a writer to use in their writing project, rather than support the reviewing process in some way. While there are exceptions [23, 29], they are in the minority [12]. Writers said the most helpful feedback was specific and they could ask questions about it. For instance, indicating a paragraph might be confusing would be less useful than indicating a particular sentence referenced an idea the reader may not have heard of before. Writers wanted to be

able to ask 'why' questions about feedback, such that they could drill into underlying issues. If a support actor says a poem is too abstract, a writer should be able to ask why and have the actor explain, perhaps, that the poem contains very few images of the natural world.[7]

*6.1.2 Explainable feedback.* The desire for feedback that is specific and can be questioned points to a potential intersection of writing support tools and explainable AI (XAI). XAI has exploded as an important field of study for any system making use of neural networks, as neural networks typically lack clear reasons for their outputs [41]. Writing feedback may prove, in addition to being an important area of study for writing tools, a useful testbed for natural language XAI systems. Writing is a complicated activity and XAI systems would have to reason about the text in a sophisticated way.

*6.1.3 AI suggestion as challenge.* Several writers brought up the idea that a computer suggestion may encourage them to come up with something better. Whereas AI writing tools are often designed to provide writers with ideas the writer would want to incorporate, and often intend to make writing easier, 'suggestion as challenge' proposes that systems may encourage writers to push themselves further, making writing more difficult but ultimately result in better writing artifacts. In this case, the rejection of a suggestion may be a sign of system success, rather than failure, and can turn the purpose of a suggestion on its head. Additionally, writers may want to put their own ideas down first, in order to preserve their pre-computational-influence ideas.

## 6.2 Research Questions and Lines of Inquiry

*6.2.1 Mental models of AI writing systems.* Since many of our participants had limited experience with an AI writing tool, they relied on preconceived notions of how AI systems tend to work. Participants with experience with SudoWrite also had varied notions of the system, and what it is and isn't good at, that may or may not align with a thorough analysis of SudoWrite's abilities. Studying the preconceived notions people have and how these notions (or mental models) change over time will be an important part of understanding how and why people interact with these systems. Studying people's mental models of AI systems is becoming a larger field of study [3, 13, 39] and exploring this issue in AI writing systems will only become more relevant as these tools are more widely adopted.

*6.2.2 Correlating attitudes with actions.* Our results point out the complicated dynamics involved with writers getting support from people or computers. These dynamics impact how writers seek out and incorporate support from people or computers. As language technologies improve, there has been more interest in studying writing support tools at scale; e.g. [22] looks at interactions with a GPT-3 autocomplete system across 63 writers. Such a line of study would benefit from also considering writers' *attitudes* towards support, such as those outlined in this paper. A survey, developed in response to the findings of this paper, would be a valuable instrument for researchers looking to understand not just how, but why, writers interact with AI writing systems in particular ways.

*6.2.3 Sharing intention with computers.* A big roadblock for any kind of support was if the support actor understood the writer's intention. Feedback often failed to address what a writer needed because the support actor didn't understand what the writer was trying to do. Similarly, writers were skeptical a computer could do this when even peers sometimes failed. However, we see a future where computers can help writers articulate or understand their own intention. While intention may not always be found in what has already been written, computers may be able to guess at the intention based what has been provided, and provide a mirror for the work that the writer may find useful. But little work today explicitly models a writer's intention, or attempts to evaluate how well a system was able to understand or align itself with an intention. More research is required to test this idea.

# 7 LIMITATIONS

## 7.1 Participant Experience with Writing Tools

While our participant population included some writers who currently use a AI writing tool, most had limited experience with such systems. This meant that their responses were colored by their preconceived notions of what such systems could and could not do. While we were purposeful with this sample, as we wanted to talk to writers from a wide range of backgrounds and writers currently using AI writing tools represent only a small slice of all writerly types, it certainly colors our results. This divide seemed to mostly impact how writers thought about their values, and if computers could live up to those values. For instance, many writers thought computers could not be creative and would only reflect 'well-trodden paths', while writers who used SudoWrite tended to think more positively about computers' creative abilities. Writers who didn't use SudoWrite tended to worry about how computational intervention would impact their feeling of authenticity, whereas SudoWrite users were less worried about this; it's unclear from our study alone if this is due an alleviation of authenticity concerns after using SudoWrite, or if SudoWrite users self-select for writers with fewer authenticity concerns.

## 7.2 Participant Expertise and Goals

Though we tried to recruit widely, our participant population limits drawing extensive conclusions. For instance, only six of our 20 participants had an Master's of Fine Art in Creative Writing; most had only informal writing education. We also didn't collect publication information, but based on the interviews it seemed that while most participants had published short pieces, many were currently attempting to publish their first book. This suggests our population may be skewed slightly toward the amateur. For instance, the two writers in our study who rarely sought out feedback were both professors, and it may be that their professional achievement gave

[7] We could imagine further follow-up questions, such as why should a poem contain images of the natural world?

them a confidence or validation that meant they were less concerned with external judgements; however they were also both poets, which also may have impacted their habits. Several other writers had great professional achievement and still relied on feedback, for instance the science journalist still made use of informal and formal external editing for feedback, and the genre writer made extensive use of alpha and beta readers. These details shed light on the complications of correlating a writer's experience with their attitudes, and more research will be needed to refine the ideas put forward in this work.

Writing genre and community may also have impacted our results. Our SudoWrite users, while perhaps representative of all SudoWrite users, were exclusively fiction writers; mostly fantasy and science fiction novelists. Their experiences with SudoWrite likely don't represent the experiences of people writing poetry, nonfiction essays, memoirs, or novels of different genres, and our inability to recruit SudoWrite users of other genres may reflect the ability of AI writing systems to cater to those types of writing. Researchers may also be interested in writing populations that may have different norms, such as fan fiction writers, where writing goals and ideas of success can be quite different than those attempting to publish their work in more traditional venues.

Finally, our writers came from an American context. Cultural norms, both around writing and technology, likely play a large role in people's attitudes towards the themes investigated in this work. Research that investigates a more global population of writers, or writers from a different cultural context, will contribute to our understanding of human-AI interaction overall.

## 8 CONCLUSION

In this work we addressed the gap between the nuanced support writers get from peers and mentors, and the more rigid and simple support writers can get from computers. We interviewed 20 creative writers from a variety of genres, including 6 writers currently using an AI writing support tool. Through a qualitative analysis, we report on three high-level categories that modulate how writers decide when and why to look for support: 1) what writers desire help with, 2) how writers perceive potential support actors, and 3) the values writers hold about the writing process. This taxonomy illuminates the social dynamics that govern writers' collaboration with outside entities, and has implications for future system design.

## ACKNOWLEDGMENTS

Katy Ilonka Gero was supported by NSF IIS 2129020.

## REFERENCES

[1] Mark S. Ackerman. 2000. The Intellectual Challenge of CSCW: The Gap Between Social Requirements and Technical Feasibility. *Human–Computer Interaction* 15, 2-3 (Sept. 2000), 179–203. https://doi.org/10.1207/S15327051HCI1523_5

[2] Stephen H. Bach, Victor Sanh, Zheng-Xin Yong, Albert Webson, Colin Raffel, Nihal V. Nayak, Abheesht Sharma, Taewoon Kim, M. Saiful Bari, Thibault Fevry, Zaid Alyafeai, Manan Dey, Andrea Santilli, Zhiqing Sun, Srulik Ben-David, Canwen Xu, Gunjan Chhablani, Han Wang, Jason Alan Fries, Maged S. Al-shaibani, Shanya Sharma, Urmish Thakker, Khalid Almubarak, Xiangru Tang, Xiangru Tang, Mike Tian-Jian Jiang, and Alexander M. Rush. 2022. PromptSource: An Integrated Development Environment and Repository for Natural Language Prompts. *arXiv:2202.01279 [cs]* (Feb. 2022). http://arxiv.org/abs/2202.01279 arXiv: 2202.01279.

[3] Gagan Bansal, Besmira Nushi, Ece Kamar, Walter S. Lasecki, Daniel S. Weld, and Eric Horvitz. 2019. Beyond Accuracy: The Role of Mental Models in Human-AI Team Performance. *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing* 7 (Oct. 2019), 2–11. https://doi.org/10.1609/hcomp.v7i1.5285

[4] Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. ACM, Virtual Event Canada, 610–623. https://doi.org/10.1145/3442188.3445922

[5] Margaret A Boden. 2009. Computer Models of Creativity. *AI Magazine* (2009), 12. https://doi.org/10.1609/aimag.v30i3.2254

[6] Kyle Booten and Katy Ilonka Gero. 2021. Poetry Machines: Eliciting Designs for Interactive Writing Tools from Poets. In *Creativity and Cognition*. ACM, Virtual Event Italy, 1–5. https://doi.org/10.1145/3450741.3466813

[7] Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. Language Models are Few-Shot Learners. *arXiv:2005.14165 [cs]* (July 2020). http://arxiv.org/abs/2005.14165 arXiv: 2005.14165.

[8] John Joon Young Chung, Shiqing He, and Eytan Adar. 2022. Artist Support Networks: Implications for Future Creativity Support Tools. In *Designing Interactive Systems Conference*. ACM, Virtual Event Australia, 232–246. https://doi.org/10.1145/3532106.3533505

[9] John Joon Young Chung, Wooseok Kim, Kang Min Yoo, Hwaran Lee, Eytan Adar, and Minsuk Chang. 2022. TaleBrush: Sketching Stories with Generative Pretrained Language Models. (2022), 19.

[10] Linda Flower and John R. Hayes. 1981. A Cognitive Process Theory of Writing. *College Composition and Communication* 32, 4 (Dec. 1981), 365. https://doi.org/10.2307/356600

[11] Samuel Gehman, Suchin Gururangan, Maarten Sap, Yejin Choi, and Noah A. Smith. 2020. RealToxicityPrompts: Evaluating Neural Toxic Degeneration in Language Models. *arXiv:2009.11462 [cs]* (Sept. 2020). http://arxiv.org/abs/2009.11462 arXiv: 2009.11462.

[12] Katy Gero, Alex Calderwood, Charlotte Li, and Lydia Chilton. 2022. A Design Space for Writing Support Tools Using a Cognitive Process Model of Writing. In *Proceedings of the First Workshop on Intelligent and Interactive Writing Assistants (In2Writing 2022)*. Association for Computational Linguistics, Dublin, Ireland, 11–24. https://doi.org/10.18653/v1/2022.in2writing-1.2

[13] Katy Ilonka Gero, Zahra Ashktorab, Casey Dugan, Qian Pan, James Johnson, Werner Geyer, Maria Ruiz, Sarah Miller, David R. Millen, Murray Campbell, Sadhana Kumaravel, and Wei Zhang. 2020. Mental Models of AI Agents in a Cooperative Game Setting. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–12. https://doi.org/10.1145/3313831.3376316

[14] Katy Ilonka Gero and Lydia B. Chilton. 2019. How a Stylistic, Machine-Generated Thesaurus Impacts a Writer's Process. In *Proceedings of the 2019 on Creativity and Cognition*. ACM, San Diego CA USA, 597–603. https://doi.org/10.1145/3325480.3326573

[15] 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*. ACM, Glasgow Scotland Uk, 1–12. https://doi.org/10.1145/3290605.3300526

[16] Katy Ilonka Gero, Vivian Liu, and Lydia Chilton. 2022. Sparks: Inspiration for Science Writing using Language Models. In *Designing Interactive Systems Conference*. ACM, Virtual Event Australia, 1002–1019. https://doi.org/10.1145/3532106.3533533

[17] John R. Hayes. 1996. A new framework for understanding cognition and affect in writing. In *The Science of Writing: Theories, Methods, Individual Differences, and Applications*. Lawrence Erbaum Associates.

[18] Julie S. Hui, Darren Gergle, and Elizabeth M. Gerber. 2018. IntroAssist: A Tool to Support Writing Introductory Help Requests. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, Montreal QC Canada, 1–13. https://doi.org/10.1145/3173574.3173596

[19] Zhengbao Jiang, Frank F. Xu, Jun Araki, and Graham Neubig. 2020. How Can We Know What Language Models Know? *Transactions of the Association for Computational Linguistics* 8 (07 2020), 423–438. https://doi.org/10.1162/tacl_a_00324 arXiv:https://direct.mit.edu/tacl/article-pdf/doi/10.1162/tacl_a_00324/1923867/tacl_a_00324.pdf

[20] R.T. Kellogg. 2008. Training writing skills: A cognitive developmental perspective. *Journal of Writing Research* 1, 1 (June 2008), 1–26. https://doi.org/10.17239/jowr-2008.01.01.1

[21] Pranav Khadpe, Ranjay Krishna, Li Fei-Fei, Jeffrey Hancock, and Michael Bernstein. 2020. Conceptual Metaphors Impact Perceptions of Human-AI Collaboration. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (Oct. 2020), 1–26. https://doi.org/10.1145/3415234 arXiv:2008.02311 [cs].

[22] Mina Lee, Percy Liang, and Qian Yang. 2022. CoAuthor: Designing a Human-AI Collaborative Writing Dataset for Exploring Language Model Capabilities. In *CHI Conference on Human Factors in Computing Systems*. ACM, New Orleans LA USA, 1–19. https://doi.org/10.1145/3491102.3502030
[23] Ming Liu, Rafael A Calvo, and Vasile Rus. 2020. Automatic Generation and Ranking of Questions for Critical Review. (2020), 15.
[24] Pengfei Liu, Weizhe Yuan, Jinlan Fu, Zhengbao Jiang, Hiroaki Hayashi, and Graham Neubig. 2021. Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing. *arXiv:2107.13586 [cs]* (July 2021). http://arxiv.org/abs/2107.13586 arXiv: 2107.13586.
[25] Mary Lou Maher. 2012. Computational and Collective Creativity: Who's Being Creative? (2012), 5.
[26] Raphaël Millière. 2022. AI Art Is Challenging the Boundaries of Curation. *Wired* (July 2022). https://www.wired.com/story/dalle-art-curation-artificial-intelligence/
[27] Meghan O'Gieblyn. 2021. Babel: Could a machine have an unconscious? (July 2021). https://www.nplusonemag.com/issue-40/essays/babel-4/
[28] Michael Quinn Patton. 2002. *Qualitative Research and Evaluation Methods (Chapter 5)* (3 ed.). Sage Publications.
[29] Zhenhui Peng, Qingyu Guo, Ka Wing Tsang, and Xiaojuan Ma. 2020. Exploring the Effects of Technological Writing Assistance for Support Providers in Online Mental Health Community. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–15. https://doi.org/10.1145/3313831.3376695
[30] Ethan Perez, Saffron Huang, Francis Song, Trevor Cai, Roman Ring, John Aslanides, Amelia Glaese, Nat McAleese, and Geoffrey Irving. 2022. Red Teaming Language Models with Language Models. *arXiv preprint arXiv:2202.03286* (2022), 31.
[31] 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
[32] Philip Quinn and Shumin Zhai. 2016. A Cost-Benefit Study of Text Entry Suggestion Interaction. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 83–88. https://doi.org/10.1145/2858036.2858305
[33] Jack W. Rae, Sebastian Borgeaud, Trevor Cai, Katie Millican, Jordan Hoffmann, Francis Song, John Aslanides, Sarah Henderson, Roman Ring, Susannah Young, Eliza Rutherford, Tom Hennigan, Jacob Menick, Albin Cassirer, Richard Powell, George van den Driessche, Lisa Anne Hendricks, Maribeth Rauh, Po-Sen Huang, Amelia Glaese, Johannes Welbl, Sumanth Dathathri, Saffron Huang, Jonathan Uesato, John Mellor, Irina Higgins, Antonia Creswell, Nat McAleese, Amy Wu, Erich Elsen, Siddhant Jayakumar, Elena Buchatskaya, David Budden, Esme Sutherland, Karen Simonyan, Michela Paganini, Laurent Sifre, Lena Martens, Xiang Lorraine Li, Adhiguna Kuncoro, Aida Nematzadeh, Elena Gribovskaya, Domenic Donato, Angeliki Lazaridou, Arthur Mensch, Jean-Baptiste Lespiau, Maria Tsimpoukelli, Nikolai Grigorev, Doug Fritz, Thibault Sottiaux, Mantas Pajarskas, Toby Pohlen, Zhitao Gong, Daniel Toyama, Cyprien de Masson d'Autume, Yujia Li, Tayfun Terzi, Vladimir Mikulik, Igor Babuschkin, Aidan Clark, Diego de Las Casas, Aurelia Guy, Chris Jones, James Bradbury, Matthew Johnson, Blake Hechtman, Laura Weidinger, Iason Gabriel, William Isaac, Ed Lockhart, Simon Osindero, Laura Rimell, Chris Dyer, Oriol Vinyals, Kareem Ayoub, Jeff Stanway, Lorrayne Bennett, Demis Hassabis, Koray Kavukcuoglu, and Geoffrey Irving. 2022. Scaling Language Models: Methods, Analysis & Insights from Training Gopher. http://arxiv.org/abs/2112.11446 arXiv:2112.11446 [cs].
[34] Nikhil Singh, Guillermo Bernal, Daria Savchenko, and Elena L. Glassman. 2022. Where to Hide a Stolen Elephant: Leaps in Creative Writing with Multimodal Machine Intelligence. *ACM Transactions on Computer-Human Interaction* (Feb. 2022), 3511599. https://doi.org/10.1145/3511599
[35] Robin Sloan. 2016. Writing with the machine. (May 2016). https://www.robinsloan.com/notes/writing-with-the-machine/
[36] Richard Jean So and Andrew Piper. 2016. How Has the MFA Changed the Contemporary Novel? *The Atlantic* (March 2016). https://www.theatlantic.com/entertainment/archive/2016/03/mfa-creative-writing/462483/
[37] Richard Jean So and Gus Wezerek. 2020. Just How White Is the Book Industry? *The New York Times* (Dec. 2020). https://www.nytimes.com/interactive/2020/12/11/opinion/culture/diversity-publishing-industry.html
[38] David R. Thomas. 2006. A General Inductive Approach for Analyzing Qualitative Evaluation Data. *American Journal of Evaluation* 27, 2 (June 2006), 237–246. https://doi.org/10.1177/1098214005283748
[39] Qiaosi Wang, Koustuv Saha, Eric Gregori, David Joyner, and Ashok Goel. 2021. Towards Mutual Theory of Mind in Human-AI Interaction: How Language Reflects What Students Perceive About a Virtual Teaching Assistant. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 384, 14 pages. https://doi.org/10.1145/3411764.3445645
[40] Tongshuang Wu, Michael Terry, and Carrie J. Cai. 2021. AI Chains: Transparent and Controllable Human-AI Interaction by Chaining Large Language Model Prompts. *arXiv:2110.01691 [cs]* (Oct. 2021). http://arxiv.org/abs/2110.01691 arXiv: 2110.01691.
[41] Feiyu Xu, Hans Uszkoreit, Yangzhou Du, Wei Fan, Dongyan Zhao, and Jun Zhu. 2019. Explainable AI: A Brief Survey on History, Research Areas, Approaches and Challenges. In *Natural Language Processing and Chinese Computing*, Jie Tang, Min-Yen Kan, Dongyan Zhao, Sujian Li, and Hongying Zan (Eds.). Springer International Publishing, Cham, 563–574.
[42] Susan Zhang, Stephen Roller, Naman Goyal, Mikel Artetxe, Moya Chen, Shuohui Chen, Christopher Dewan, Mona Diab, Xian Li, Xi Victoria Lin, Todor Mihaylov, Myle Ott, Sam Shleifer, Kurt Shuster, Daniel Simig, Punit Singh Koura, Anjali Sridhar, Tianlu Wang, and Luke Zettlemoyer. 2022. OPT: Open Pre-trained Transformer Language Models. http://arxiv.org/abs/2205.01068 arXiv:2205.01068 [cs].

# A METHODS

## A.1 SudoWrite Features

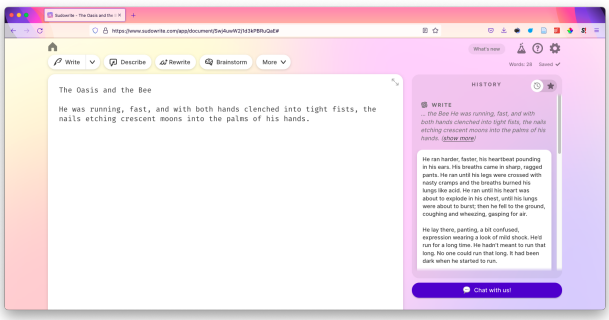


**Figure 6: 'Write' feature of SudoWrite, which continues from wherever the cursor is.**

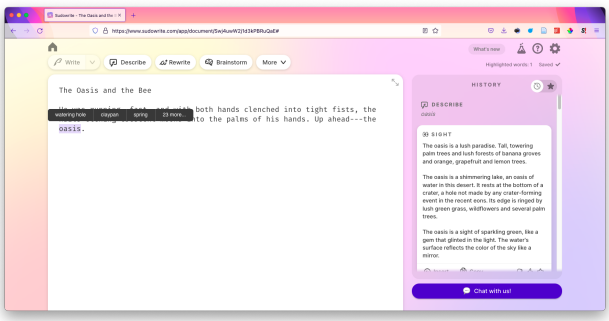


**Figure 7: 'Describe' feature of SudoWrite, which generates descriptions for highlight words or phrases.**

Figures 6, 7, and 8 demonstrate some of the functionality of SudoWrite at the time of this study. Suggested text is shown on a panel on the right, where multiple options are available. For instance, when using the 'write' functionality, multiple different continuations are generated, and writers can easily paste them into the main text box. For the 'describe' functionality, different kinds of descriptions are generated, such as 'sight', 'smell', and 'metaphor'. The 'brainstorm' functionality has its own interface separate from the writing interface.

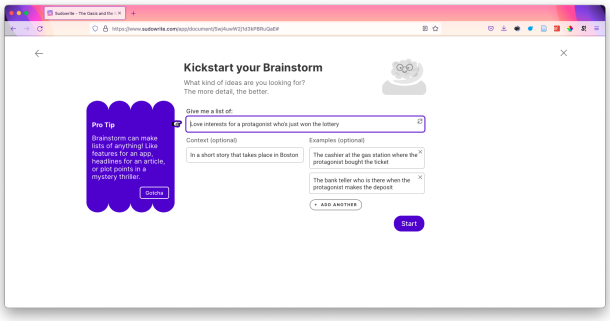


**Figure 8: 'Brainstorm' feature of SudoWrite, which generates ideas based on some context.**

SudoWrite functionality has changed over time. Some users talked about the 'wormhole' function, which seemed to be an earlier version of the 'write' function that is no longer available. These screenshots are intended to reflect SudoWrite functionality at the time the interviews were done.

## A.2 Interview Questions

■ **General questions about writing**

(1) What kind of writing do you currently do? Have done in the past?
(2) How long have you been writing?
(3) What kind of formal or informal education have you had as a writer?
(4) What is a piece of writing you are very proud of? Why?
(5) Walk me through the process or life cycle for the last piece you wrote.
Be sure to ask about external influences like research and feedback.
(6) Do you ever get writer's block? What do you do?
(7) Do you ever feel stuck revising something? What do you do?

■ **Questions about existing influence**

(8) Are there people who currently influence or in the past have influenced your writing? Who? In what ways?
Does this happen abstractly (this teacher had a big impact on me) or more concretely (this teacher influences me when they give me feedback).
(9) Are there texts or writers — who you haven't interacted with personally — who currently influence or in the past influenced your writing? Who/what? In what ways?
Again, does this happen abstractly or concretely?
(10) Someone giving you an assignment?
(11) How about editors?
(12) Do you use dictionaries, thesauruses, or other kinds of references when you write? Which? In what ways?
(13) Have you done collaborative writing projects? Please share details.
(14) Do you like to try new writing forms or styles? Why or why not?
(15) Do you like to seek out feedback? Why or why not?

■ **Questions about computer influence**

Question template: If a computer program could _____ like a teacher or peer could, would you use it? Why or why not?
(16) suggest places to revise
(17) rewrite sections
(18) write in a gap
(19) finish a piece
(20) continue something when you felt stuck
(21) something about reader perspective – where a reader might get stuck or bored or confused
(22) write into something out of domain – explanation or science fiction details
(23) do research for you or describe a city or summarize a new technology

■ **Sudowrite user questions**

(Replaces 'Questions about computer influence')
(24) What attracted you to SudoWrite?
(25) Could you talk about a specific moment you used SudoWrite and what it looked like?
(26) Which features do you use the most, and why?
(27) In what ways do you feel like SudoWrite is 'human-like' or not in its capabilities?
(28) Are there parts of your writing process you would not use SudoWrite (or something similar) for?
(29) How do you feel SudoWrite impacts your writing?
For example, how would your writing be different if you didn't use SudoWrite?
(30) Does SudoWrite perform functions that otherwise might be performed by a peer, mentor, or editor?