A Design Space for Writing Support Tools Using a Cognitive Process Model of Writing

Katy Ilonka Gero
Columbia University
katy@cs.columbia.edu

Alex Calderwood
University Of California, Santa Cruz
alexcwd@ucsc.edu

Charlotte Li
Columbia University
li.zihao@columbia.edu

Lydia B. Chilton
Columbia University
chilton@cs.columbia.edu

Abstract

Improvements in language technology have led to an increasing interest in writing support tools. In this paper we propose a design space for such tools based on a cognitive process model of writing. We conduct a systematic review of recent computer science papers that present and/or study such tools, analyzing 30 papers from the last five years using the design space. Tools are plotted according to three distinct cognitive processes—planning, translating, and reviewing—and the level of constraint each process entails. Analyzing recent work with the design space shows that highly constrained planning and reviewing are under-studied areas that recent technology improvements may now be able to serve. Finally, we propose shared evaluation methodologies and tasks that may help the field mature.

1 Introduction

The development of large-scale language models (sometimes called foundation models) is dramatically changing what technology can achieve and support (Bommasani et al., 2021). Language models like GPT3 (Brown et al., 2020) and Meena (Adiwardana et al., 2020) have led to an increasing interest in how these new technologies may support writers, for instance by providing a journalist with text in the style of The New Yorker (Seabrook, 2019) or giving a novelist a new story ending (Marche, 2021). In this paper we seek to understand where research on writing support tools currently stands, and what areas of research may be important but currently under-served.

Computational approaches to writing support have a long and rich history, certainly dating back to before the introduction of modern computation, at least to the early 1900s with the cut-up method (Burroughs, 1961) and 'plot genie' books (Hill, 1931), and likely even further back when considering the long history of generative traditions such as tarot cards (Sullivan et al., 2018). In more contemporary understandings of computation, technology developed by the natural language processing (NLP) community is often taken up as a writing tool.[1] We believe the advent of foundation models poses an exciting inflection point at which these technologies can be used to support the evergreen activity of writing in new ways.

In this paper, we draw on a cognitive process model of writing that considers writing to be a goal-directed thinking process with three distinct and non-linear cognitive processes: planning, translating, and reviewing (Flower and Hayes, 1981). We use this model to propose a design space for writing support tools. This allows us to understand what a writing support tool is attempting to support, and identify gaps or opportunities in the field. It provides a shared vocabulary for researchers, and we hope it will help the field mature and provide common goals and methodologies.

To demonstrate the use of the design space, we perform a systematic literature review of research on writing support tools from the last five years (2017-2021). This shows areas of active research and under-served areas, as well as limitations of current technology to support different aspects of writing. We also use these papers to investigate how to evaluate writing support tools.

The contributions of this paper are:

- A design space for writing support tools, based on a cognitive process model of writing.

- A systematic literature review of writing support tools ($n_{papers} = 30$) from 2017-2021.

- A gap analysis highlighting opportunities for designing future writing support tools.

- A series of common evaluation methodologies for future work to draw on.

---

[1] For example, spell-checking was an early use of pointwise mutual information (Peterson, 1980), the exciting NLP technology of its time.
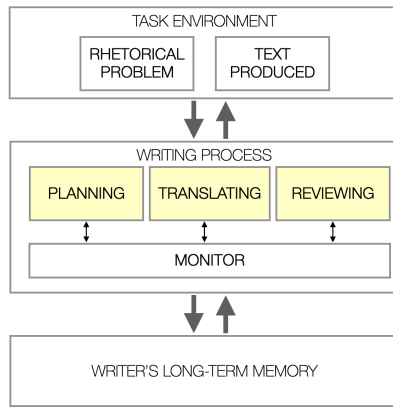
Figure 1: The cognitive process model for writing, as proposed by Flower and Hayes (1981).

2 Related Work

2.1 A Cognitive Process Model of Writing

Flower and Hayes (1981) theory of the cognitive processes involved in writing laid the groundwork for a plethora of research on the psychology of writing over the past four decades. This process model, backed by empirical studies, proposed that writing is best understood as a set of distinct hierarchical thinking processes. Figure 1 shows a schematic of the model, with the three main writing processes—planning, translating[2], and reviewing—highlighted in yellow. When Flower and Hayes state that these processes are hierarchical, they mean they can be called upon iteratively, being embedded within each other. For example, when a writer is constructing a sentence (translating), they may call in a compressed version of the entire writing process. Flower and Hayes' are also quick to note that these processes are not linear. While a common mantra is to 'plan, then write, then review', in reality writers are making plans and reviewing what they have written all throughout the writing process.

Flower and Hayes also proposed that the act of writing is propelled by goals, which are created by the writer and grow in number as the writing progresses. These goals, which span in complexity and abstraction from 'appeal to a broad audience' to 'don't use that cliche', are what direct the writer to different processes. We can model the writing process by considering the writer's goals and what

processes they enlist to achieve them.

While this model has since been updated with an increase in complexity[3], considering how goals propel the writing process remains a useful model. Writing has long been considered a mode of learning, as it is both a process and a product, which allows near-constant reflection on the ideas the writer is trying to express (Emig, 1977). By considering a writer's shifting goals, writing researchers have understood why mature writers are able to learn from their writing (Scardamalia and Bereiter, 1987).

We make use of this theory to structure a design space for writing support tools: to understand what these tool actually help with, and how we might design new ones. While there are many ways to think about writing and how computers may support it, we focus on the cognitive process model as it emphasizes writers' intentions, rather than their actions. We believe that this abstraction away from the mechanics of writing will help researchers articulate their intentions with writing support tools, and share results across disparate writing tasks.

2.2 Design Spaces

One way to synthesize a multitude of designs is to envision it in a 'design space', or a metaphorical laying out of designs according to some metrics or measures. MacLean et al. (1996) describe design space *analysis* as an approach to representing design rationale. In this view, a design space places a design in a "space of possibilities" and uses this placement to explain why a design was chosen among all the various possibilities. This frames design spaces as a useful way of communicating with stakeholders. By explaining why a design was chosen, stakeholders can better sell, maintain, and otherwise interact with a product.

Woodbury and Burrow (2006), addressing the growing popularity of design spaces in computational research, describe design space *exploration* as the idea that we can use exploring alternatives as a compelling model of design. This involves representing designs in a meaningful way, and using the representation to explore the space.[4]

A popular and highly-cited example of a design space comes from wireless sensor networks (Romer and Mattern, 2004). As the use of such networks

---

[2]They use the word 'translating' to refer to the act of putting words on the page, as 'writing' is used to describe the whole process and 'composing' can also be ambiguous. While 'translating' is typically used in NLP communities to denote converting text between languages, we use it here as a technical term to aligns with relevant psychology research.

[3]Hayes adds much more detail to the long-term memory component, and adds components for working memory and the motivation and affect of the writer (Hayes, 1996)

[4]It can also be used to build computer systems to aid in the exploration.

increased globally, "it was very difficult to discuss specific application requirements, research directions, and challenges." The proposed solution was a sensor network design space: its various dimensions would be categorized in order to both understand the existing research as well as discover new designs and applications. One conclusion was that a small set of platforms could cover the majority of the design space, rather than requiring numerous, application-specific platforms.

In this paper we introduce a design space both to think about what writing support tools currently do, and what we might want them to do in the future. In this sense we take both MacLean's and Woodbury's view: the design space is both a way to talk about why existing tools are the way they are, as well as a way to design new ones.

2.3 Related Literature Reviews

Related work has looked at a design space for non-visual word completion (Nicolau et al., 2019) and hybrid paper-digital interfaces (Han et al., 2021). We look to these for methodologies and areas of overlapping interest. Perhaps more related is work from Strobl et al. (2019) in which they perform a review on digital support for academic writing. They review 44 papers addressing essay writing needs in US secondary school instruction. Many of these papers come from educational research communities, and few use NLP technologies. Our review focuses more on human-computer interaction communities and leans more towards system that incorporate NLP technologies. When performing our literature review, we follow the checklist outlined in PRISMA[5] for performing a systematic literature review, including specifying inclusion / exclusion criteria and all sources searched.

3 Writing Goals Design Space

Flower and Hayes (1981) describe writing in the following way:

> The act of composing itself is a *goal-directed thinking process*, guided by the writer's own growing network of goals.

These writing goals may be large, like to write up an experiment for an academic paper, or small, like to make a sentence sound more formal. They may be open-ended, like to come up with the name for a

character, or quite limited, like to spell a word correctly. The goals may require imagining the reader, like to determine if a sentence is too confusing, or they may require diving deeper into what's already written, like to ensure a technical topic is discussed consistently throughout an article. Writing goals may start as external motivators—someone may ask one to write something—but as one writes, writing goals are created by the writer and propel the writing process forward.

We propose using this to structure a design space for writing support tools. Whether we call them support tools, assistants, co-creators or machines-in-the-loop, we believe what unites these systems is that they take on goals inherent in the writing process. We propose two axes for the design space:

1. Which part of the writing process the system aims to support. Flower and Hayes, in their original model of writing, propose three components: planning, translating, and reviewing. These three components align with models of creativity, which often cite ideation, implementation, and evaluation (Amabile, 1983). In both cases the components are accessed iteratively, and often hierarchically. A writer may start with a high-level plan, and then in the act of 'translating' the plan may create a smaller plan within it. Splitting up writing support tools into these processes helps us understand how, when, and why a writer may use a tool.

We acknowledge that there can be some ambiguity in distinguishing between these processes. For instance, consider a tool that, upon request, completes a writer's sentence. This tool may be supporting translating, if the completion is intended to articulate what the writer already had in mind. Or it could be supporting planning, if the completion is intended to provide the writer with new ideas or directions for their writing. When annotating papers, we rely on how the researchers describe the tool, though we acknowledge the ambiguities involved in this and that writers may use a tool in unexpected or unintended ways.[6]

2. The amount of constraint the goal has. A highly constrained goal has very few possible solutions, like when writing a technical definition. A lightly constrained goal has many possible solutions, like when describing a newly introduced fictional character. The amount of constraint gives

---

[6]An alternate approach is to rely on how writers describe their usage, but given that many papers did not include this in their evaluation, we would not have been able to annotate all papers using this method.

```
plan
open-ended                    historical fiction plot
plot ideation                              ideation

translate
open-ended                              automatic
story generation                        summarization

review
high-level feedback            suggesting stylistic
                                         revisions

              level of constraint  →
```
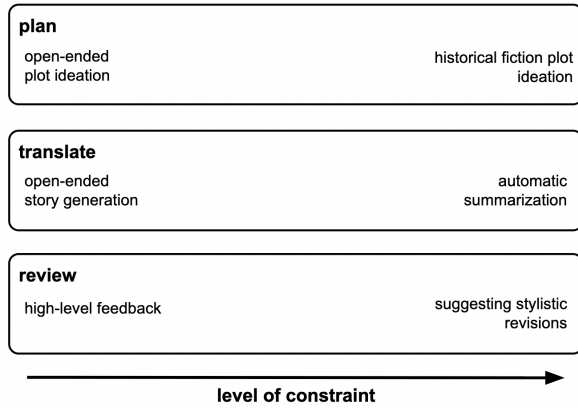
Figure 2: The writing goals design space is defined by the part of the writing process a tool wants to support and the level of constraint of the goal. This shows some example writing goals a tool may want to support.

us a measure of how particular the support must be to achieve the goal. This may be considered a measure of difficulty—writing a technical definition is very constrained, and supporting this writing task requires a high level of world understanding from a system—but constraint doesn't always imply difficulty. A writing goal may be very constrained, for instance make a particular sentence more positive, but the support may be fairly straightforward, like providing a list of positive words.

Figure 2 shows some hypothetical writing support tools in this design space, to better understand the space. Further details and descriptions of the design space can be found in the Appendix.

4 Methodology

We perform a preliminary, systematic literature review such that we can plot tools in the design space. This validates the utility of the design space and provides insights into the landscape of writing support tools.

4.1 Designing a Search Query

We design a query for searching the ACM Digital Library for relevant papers. Our goal for this query is to find as many relevant papers as possible, while minimizing the number of irrelevant papers needed to sort through. This proved more difficult than expected because search terms like 'writing' and 'support' are quite common in other subfields, like those studying memory architecture. We iterated on a query that returned many of the papers we expected to be included (such as (Roemmele and Gordon, 2018a) and (Wambsganss et al., 2020)),

while also returning less than 300 results, such that we could visually inspect them all in a timely manner. We chose to only look at papers from the last five years as we wanted to focus on where the field is currently going. We didn't require an average yearly download or number of citations, as done in other systematic reviews like Frich et al. (2019), because we wanted to include very recent work that may not be well-distributed yet.

Our final query can be found in the Appendix. It resulted in 216 items.

4.2 Selecting Papers to Include in Review

First we had one researcher read the titles of all papers and perform a quick 'desk reject' on any papers that were clearly off topic.[7] After this, 77 papers remained. Of these papers, two researchers read all the abstracts and noted if they thought a paper should be included based on the inclusion criteria below. They did this separately, and then came together to discuss and resolve disagreements.

Our inclusion criteria was:

1. a conference or journal publication[8]

2. a contribution that presents or studies a tool that aids in the translation of ideas into text

We include additional examples of what would and would not be included (which the researchers used as guidelines) in the Appendix.

This resulted in 30 papers. A list of these papers can be found in the Appendix. Each paper was assigned a nickname which allowed for easier reference than the paper title or author list.

4.3 Annotating the Selected Papers

Three members of the research team participated in the annotations. The selected papers were split up, and each paper was annotated by a single researchers. Some of these annotations were to allow us to plot tools in the design space, others were to align with Frich et al. (2019), a systematic review of creativity support tools, and still others were to quantify the type of contribution. The full list of annotations, as well as details on how ambiguities in the annotations were resolved, can be found in the Appendix. The results of our annotations can be found at `https://github.com/kgero/writing-support-tools-2022`.

---

[7]For example, a paper with the title 'A Tool for Visualizing Classic Concurrency Problems' was rejected for clearly being about a different topic.

[8]i.e. not a course description, workshop proceedings, etc.

5 Results and Analysis

5.1 The Writing Goals Design Space

In this section we consider how tools are distributed in the design space, which looks at the type of goal the tool supported, and how constrained that goal is. The 30 papers represented 33 systems, with some papers presenting multiple systems.[9] Three papers studied tools that supported all parts of the writing process: Writing Together (Olson et al., 2017) studied Google Docs, Writing on Github (Pe-Than et al., 2018) studied GitHub, and Literary Style (Sterman et al., 2020) presented an early stage exploratory tool. We exclude these because it is difficult to locate them in a single part of the design space; future work may consider how tools can be distributed across multiple parts of the design space. Excluding these, we are left with 27 systems to analyze in this section.

Figure 3 shows all tools in the writing goals design space. We color them by the size of the goal being supported. We see most parts of the design space covered, with tools in all three parts of the writing process and spanning many different levels of constraint. The papers also operate on all different sizes of writing goals.[10]

The design space shows that planning and reviewing lack work on highly constrained support, suggesting an area for future work. As the constraint for the goal increases, tools tend to support narrower and more structured writing tasks. In planning, MiL (stories) (Clark et al., 2018) and BunCho (Osone et al., 2021) (both constraint=1) support any kind of story writing, while MiL (slogans) (Clark et al., 2018) and Metaphoria (Gero and Chilton, 2019b) (both constraint=4) support slogan and metaphor writing, which have rules and syntactic structures to guide the generation process. Reviewing similarly sees this move towards the niche as constraint increases. Textlets (Han et al., 2020) (constraint=1) is a general purpose reviewing tool based on a sophisticated usage of the 'find' command. In contrast, MepsBot (Peng et al., 2020) (constraint=4) focuses on comments in online mental health forums and Dajke (Schmidt, 2020) (constraint=5) is about adjusting the reading

level of Tibetan learning material. Lightly constrained support for planning often relies on newer text generation technologies: MiL (stories) (constraint=1) and MiL (slogans) (constraint=4) come from the same paper (Clark et al., 2018), but the lightly constrained work on stories relies on a neural network while the more constrained work on slogans relies on templates.

Does a highly constrained writing goal need to be niche or highly structured? It may be that language technologies have not yet been capable of supporting more general purpose but still highly constrained writing goals. For instance, brainstorming often happens at multiple points throughout a creative process, with later brainstorming being more constrained by previous choices. Early stage brainstorming may be easier to support because there are less constraints needed to get right. An area new technologies could explore is later-stage brainstorming, which could be quite general purpose—input any piece of writing and a brainstorming prompt—but still lie in the highly constrained planning part of the design space.

The design space shows that highly constrained support for translation is well studied; these systems tend to support highly structured writing tasks. AmbientLetter (Toyozaki and Watanabe, 2018) supports spell-checking while writing on paper; LyriSys (Watanabe et al., 2017) generates topically relevant song lyrics based on a syllabic pattern; Play Write (Iqbal et al., 2018) supports writing microtasks; StoryAssembler (Garbe et al., 2019) supports writing dynamic / non-linear stories. Because the writing goals are quite diverse, these systems use a variety of technologies. Some are about providing text to the writer but most provide support in some other way, like structuring tasks or ensuring constraints are met.

As in planning and reviewing, the translating tools for highly constrained goals are more highly structured. Likely this structure is what allows the tool to be supportive, or is developed by designers to provide traction for the problem. We also saw these tools being quite niche. More general writing tasks like storytelling (e.g. MiL (stories) (Clark et al., 2018), BunCho (Osone et al., 2021), and Writing with RNN (Roemmele and Gordon, 2018b)) were lightly constrained, but this isn't inherent to storytelling. Subtasks within storytelling can be quite constrained, but we didn't see them turn up in our literature review. An interesting

---

[9] UI Design (Gonçalves and Campos, 2017) studied four systems, but since they were all very similar, for this section we consider them to be a single system (as they would be in the same part of the design space anyway).

[10] 5 at the level of words, 6 at sentences, 8 at paragraphs, 3 at more than the paragraph, and 5 on the writing experience.
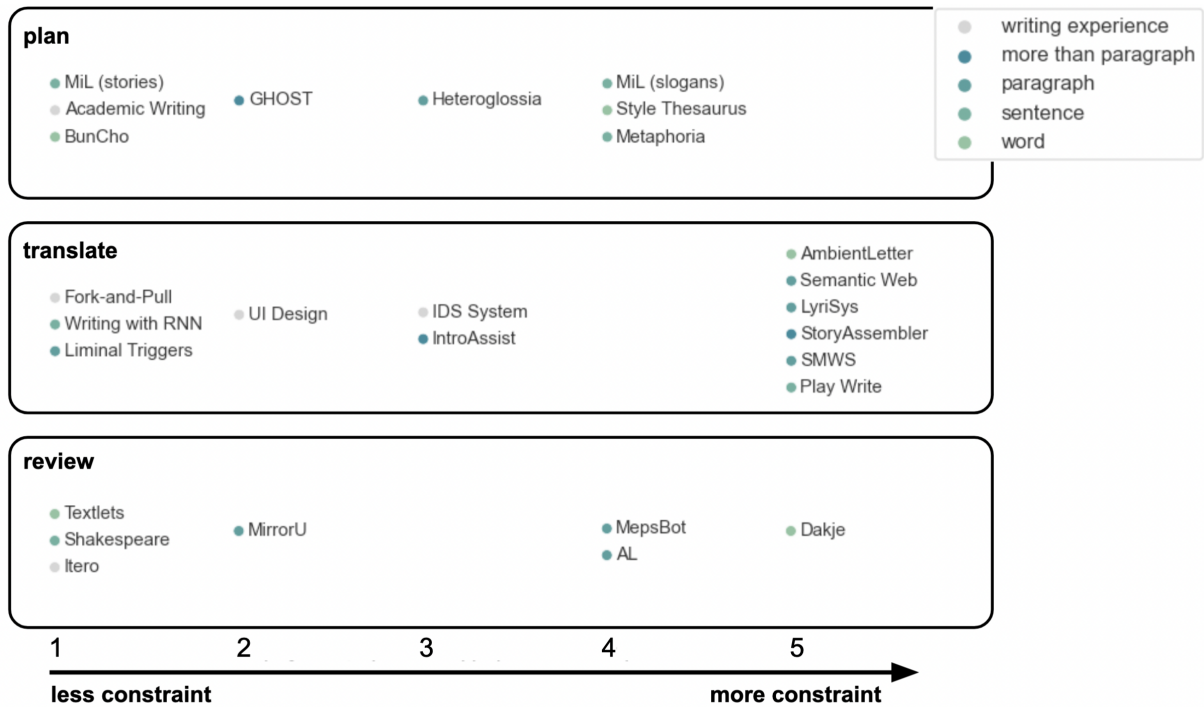
Figure 3: Twenty-seven writing support tools plotted in the writing goals design space. We can see that highly constrained planning and reviewing are under-explored areas.
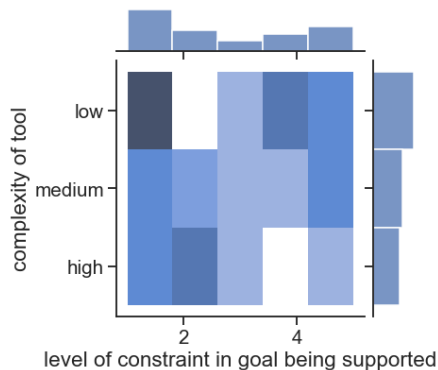

Figure 4: There were more tools with 1-2 features (low complexity). The distribution of constraints being supported was U-shaped.

example of highly constrained translation that we didn't see is taking bullet points and turning them into prose. This is another example of a highly constrained but more general purpose task we believe is an interesting area for future work.

5.2 Complexity of Tool and Technology Used

The tools studied had various levels of technical complexity, drawing on a wide spectrum of user interactions and language technologies. They ranged from full document editors such as Microsoft Word and OmniFocus, which provide rich interface's on top of feedback such as spell checking, to collabo-

ration software such as GitHub, to text generation technologies such as context-free grammars and neural algorithms. Figure 4 shows the distribution of tools according to complexity and level of constraint. For annotating the complexity of a tool we followed Frich et al. (2019), where high complexity refers to an entire system or suite of tools, and low complexity refers to tools with only one or two features. (That is, complexity here is not a measure of technical difficulty.) The tools reviewed were slightly skewed towards low complexity (14 of the 33 tools). Most of the tools (78%) were contributions of the authors.

A third (11 of 33) of the tools used a neural algorithm for text generation or translation and five used some other form of grammar, template, or external knowledge source for text generation. BunCho (Osone et al., 2021) was one of the handful of non-English tools (5 of 33), using GPT-2 to generate Japanese story titles and summaries. Predictive text completion was used by a number of tools, like Storytelling Assistance (Roemmele and Gordon, 2018a), to insert text in a way that might provoke the writer to explore new directions and see their work in a new light.

A number of the tools were more highly constrained, providing some form of scaffold or guidance. Tools like IntroAssist (Hui et al., 2018) use

cognitive writing theories to produce static scaffolds that assist writers in their goals, in this case to write an intro email. Style Thesaurus (Gero and Chilton, 2019a) and Metaphoria (Gero and Chilton, 2019b) were among the more highly constrained tools that served as ideation support; the latter generating metaphors from input terms rather than producing sentence-level text.

A number of the tools were interested in analyzing and improving written text at various intermediate points in the writing process. Itero (Türkay et al., 2018) visualized document revision statistics to let writers get a better sense of their own interaction with their written words. AL (Wambsganss et al., 2020) used natural language processing to provide feedback on the quality of essays in terms of their argument structure, readableness, and coherence. Of these, some went the further step of correcting or altering the writer's text. SMWS (Wu et al., 2019) used the paradigm of neural text translation to 'translate' a Dyslexic writer's Facebook comments into non-Dyslexia style writing.

The front-end user experience was primary to many of the tools. UI Design (Gonçalves and Campos, 2017) investigated how various interfaces promoted focus and other such writing considerations, and which led to increased writing quality. Liminal Triggers (Gonçalves et al., 2017) built an editor to investigate the effectiveness of subliminal priming to reduce writer's block. Textlets (Han et al., 2020) turned selected text into manipulable objects for intradocument organization. A few of the studies were interested in situating writing interfaces into alternative environments, such as a smartphone app for mixed-attention environments (Iqbal et al., 2018) and game-text writing tool embedded right into the game engine (Guarneri et al., 2017).

Many of the tools employed networking. Writing Together (Olson et al., 2017) examined the collaborative effects of Google Docs, a full web-based writing interface with inline comments and tracked revision histories. IDS (Tian et al., 2021) provided a mechanism to collaboratively turn summary writing into the form of a final document. A few of the studies explored how GitHub's pull/push workflow, which differs subtitantively from the live-editing affordances of Google Docs, can be used to improve writing quality. Heteroglossia (Huang et al., 2020) expands the typical idea of collaboration with a system that had Mechanical Turkers roleplay for individual characters within a creative story.

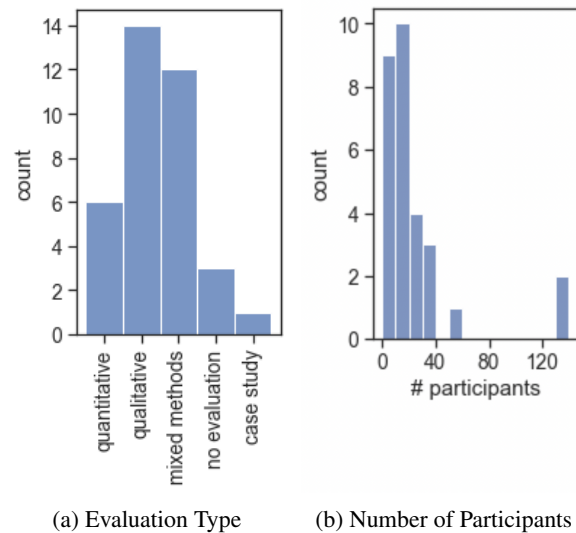(a) Evaluation Type　　(b) Number of Participants

Figure 5: Histograms representing the distribution of evaluation methodologies.

5.3 Analysis of Evaluation Methodologies

A total of 33 evaluations were conducted among the 30 papers we studied. Several papers conducted more than one evaluation for their research, while three papers had no evaluation: Shakespeare (Liu et al., 2019), Dakje (Schmidt, 2020), and Ambient Letter (Toyozaki and Watanabe, 2018).

Figure 5 shows the distributions of evaluation type and number of participants. On average, 25 participants were recruited for evaluation of writing tasks. 75% of the evaluations were conducted with fewer than 40 participants and these evaluations were either qualitative or mixed methods, likely because qualitative evaluations produce large and unorganized data that does not allow easy manipulation and analysis for too many participants. Writing Together (Olson et al., 2017) and Storytelling Assistance (Roemmele and Gordon, 2018a) conducted studies with about 130 participants, and both were quantitative only evaluations.

Looking at the papers that had some component of qualitative evaluation, there was a wide range of criteria studied, including quality of writing, usability, usefulness, coherence to context, enjoyment, satisfaction, impact on flow, impact on confidence, and many more. Qualitative studies tended to assess their tools through semi-structured interviews with a small group of target users, such as creative writers or students. Around 50% of qualitative evaluations were done alongside a quantitative evaluation. Studies with only quantitative evaluations, such as Storytelling Assistance (Roemmele and

Gordon, 2018a), assessed quality of the tool with questionnaires reported on a Likert scale and used measures specific to the tools they are studying, like Levenshtein edit distance or simultaneous time spent on writing, to evaluate user's attitudes and collaborative usage of the tool.

Around half of the evaluations reported did not include the time participants spent writing with the system, which makes it difficult to assess this in relation to other aspects of the studies. Among the evaluations that reported time spent writing, quantitative evaluations done without the addition of a qualitative evaluation have a much shorter average time spent with the user (5-10 mins) than the others (25 mins). However, there's nothing inherent about quantitative or larger-scale evaluations that precludes writing for a longer period of time.

Quality of writing corresponds to a variety of different task-specific measures. MiL (stories) (Clark et al., 2018) has Amazon Mechanical Turk workers rate outputs for creativity, coherence, grammaticality, and entertainment. AL has annotators rate an argument according to a formal schema. Writing Together (Olson et al., 2017) studied writing done during a project writing course; writing quality was determined by course graders.

Given so much variety in the evaluation methodologies, we make several recommendations on how evaluations could become more comparable:

- Report more details of the actual writing done in the study, for instance amount of time spent writing, amount of words written, and the type of participants recruited (novice, expert, etc.).
- Use shared surveys rather than develop new ones each time. The Creativity Support Index (Cherry and Latulipe, 2014), NASA Task Load Index (Hart and Staveland, 1988), and Technology Acceptance Model (Venkatesh and Davis, 2000) may all be useful. We also encourage researchers to propose writing-specific surveys that can be used by others.
- Report user interaction measures, like edit distance, and number and frequency of interactions, that can be shared across different writing tasks.

Perhaps the biggest barrier for comparing research is the lack of shared tasks. These papers represent a broad range of writing tasks, from slogan writing to dynamic storytelling to argumentative writing. While we do not believe that writing is a monolith, and nor should be writing support tools, a set of shared tasks may help consolidate the work.

We suggest three shared writing tasks: story writing (fiction), argumentative essay writing (nonfiction), and personal essay writing (creative nonfiction). Personal essay writing has many elements of fiction, like relying on character and narrative, while being constrained to the reality of the writer's lived experience. These tasks span from being completely open-ended (story writing) to partially constrained (personal essay) to quite constrained (argumentative essays). Within each task are many subtasks which span from being very open-ended (how to start the argumentative essay) to very constrained (how to describe an existing character).

We choose these tasks because they each contain goals which could span the entire design space and a variety of genres. There are many tasks we did not include, like emails, explainers, and poetry. These were not chosen because we felt they were too niche (like poetry) or too broad-reaching (like emails) to help unify research.

Below we discuss some variation within each task, and some potential subtasks to focus on:

- Story writing. This already-common task contains within it diverse goals from plot development to scene description. The length can vary its complexity and they can be constrained to varying degrees by a prompt. We recommend two specific tasks. The first is writing stories in response to a prompt. (Again, this is already common and can be continued to be worked on.) The second is adding detail to an existing or partially written story, for instance adding character or scene descriptions. This will allow work to look at some of the more constrained parts of story writing.
- Argumentative essay writing. This task is common in U.S. secondary education and can be extended to include journalistic forms like opinion pieces. It contains subtasks like defending propositions, writing an engaging introduction, and appealing to the audience. We recommend two specific avenues of research: Supporting argumentative structure, and supporting introductory remarks. While supporting structure gets to complicated technical elements of the ideas of a piece of writing, supporting introductory remarks requires more modeling of the reader and understanding what makes text interesting and engaging.

- Personal essay writing. This task can include private journaling as well as more public forms like memoir or even personal statements. It can contain subtasks like finding relevant historical information or identifying potential narratives. The utility of this task is how writers are self-motivated. For this task we recommend focusing less on the quality of writing, and more on the experience of the writer. While stories and argumentative essays have many formal elements that can be used in evaluation, we recommend this task be about immersion and self-expression.

6 Limitations

Our systematic review was limited in scope, as we focused only on the last five years, and our query for selecting papers may not have caught all relevant papers. For instance, one clear problem with using the ACM Digital Library is that many NLP conferences are not included. Future work should investigate more sources for papers, and look further into the archive. Additionally, we did not include commercial or open source writing tools that exist outside of the academy, which likely would improve the findings of any large-scale, systematic review of writing support tools.

There are also many more questions that could be asked about writing support tools. For instance, we found that user type was not widely reported, but user type may be implied by the writing task, or inferred by the evaluation methodology. Relatedly, further analysis could be done on how much work is dedicated to fiction v. nonfiction or short v. longer writing. We hope that by making our selected papers easily accessible, others may use this to do their own investigations with other focuses.

7 Conclusion

We present a design space for writing support tools based on a cognitive process model of writing. We perform a systematic literature review, reviewing 30 papers from the last five years (2017-2021). We find that highly constrained planning and reviewing are under-studied areas. We see that evaluation methodologies vary widely, and propose validated surveys and interaction measures as ways to make evaluations more comparable across systems. We also propose three shared tasks—storytelling, argumentative writing, and personal essays—to aid in propelling work on writing support tools forward.

References

Daniel Adiwardana, Minh-Thang Luong, David R. So, Jamie Hall, Noah Fiedel, Romal Thoppilan, Zi Yang, Apoorv Kulshreshtha, Gaurav Nemade, Yifeng Lu, and Quoc V. Le. 2020. Towards a Human-like Open-Domain Chatbot. *arXiv:2001.09977 [cs, stat]*. ArXiv: 2001.09977.

Teresa M Amabile. 1983. The social psychology of creativity: A componential conceptualization. *Journal of personality and social psychology*, 45(2):357.

Rishi Bommasani, Drew A Hudson, Ehsan Adeli, Russ Altman, Simran Arora, Sydney von Arx, Michael S Bernstein, Jeannette Bohg, Antoine Bosselut, Emma Brunskill, et al. 2021. On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258*.

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]*. ArXiv: 2005.14165.

William S Burroughs. 1961. The cut-up method of brion gysin. *The third mind*, pages 29–33.

Erin Cherry and Celine Latulipe. 2014. Quantifying the Creativity Support of Digital Tools through the Creativity Support Index. *ACM Transactions on Computer-Human Interaction*, 21(4):1–25.

Elizabeth Clark, Anne Spencer Ross, Chenhao Tan, Yangfeng Ji, and Noah A. Smith. 2018. Creative writing with a machine in the loop: Case studies on slogans and stories. In *23rd International Conference on Intelligent User Interfaces*, IUI '18, page 329–340, New York, NY, USA. Association for Computing Machinery.

Janet Emig. 1977. Writing as a Mode of Learning. *College Composition and Communication*, 28(2):122–128. Publisher: National Council of Teachers of English.

Linda Flower and John R. Hayes. 1981. A Cognitive Process Theory of Writing. *College Composition and Communication*, 32(4):365.

Jonas Frich, Lindsay MacDonald Vermeulen, Christian Remy, Michael Mose Biskjaer, and Peter Dalsgaard. 2019. Mapping the Landscape of Creativity Support Tools in HCI. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–18, Glasgow Scotland Uk. ACM.

Jacob Garbe, Max Kreminski, Ben Samuel, Noah Wardrip-Fruin, and Michael Mateas. 2019. Storyassembler: An engine for generating dynamic choice-driven narratives. In *Proceedings of the 14th International Conference on the Foundations of Digital Games*, FDG '19, New York, NY, USA. Association for Computing Machinery.

Katy Ilonka Gero and Lydia B. Chilton. 2019a. How a stylistic, machine-generated thesaurus impacts a writer's process. In *Proceedings of the 2019 on Creativity and Cognition*, C&C '19, page 597–603, New York, NY, USA. Association for Computing Machinery.

Katy Ilonka Gero and Lydia B. Chilton. 2019b. Metaphoria: An algorithmic companion for metaphor creation. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, page 1–12, New York, NY, USA. Association for Computing Machinery.

Frederica Gonçalves and Pedro Campos. 2017. Understanding and evaluating the user interface design for creative writing. In *Proceedings of the European Conference on Cognitive Ergonomics 2017*, ECCE 2017, page 85–92, New York, NY, USA. Association for Computing Machinery.

Frederica Gonçalves, Ana Caraban, Evangelos Karapanos, and Pedro Campos. 2017. What shall i write next? subliminal and supraliminal priming as triggers for creative writing. In *Proceedings of the European Conference on Cognitive Ergonomics 2017*, ECCE 2017, page 77–84, New York, NY, USA. Association for Computing Machinery.

Andrea Guarneri, Laura A. Ripamonti, Francesco Tissoni, Marco Trubian, Dario Maggiorini, and Davide Gadia. 2017. Ghost: A ghost story-writer. In *Proceedings of the 12th Biannual Conference on Italian SIGCHI Chapter*, CHItaly '17, New York, NY, USA. Association for Computing Machinery.

Feng Han, Yifei Cheng, Megan Strachan, and Xiaojuan Ma. 2021. Hybrid Paper-Digital Interfaces: A Systematic Literature Review. In *Designing Interactive Systems Conference 2021*, pages 1087–1100, Virtual Event USA. ACM.

Han L. Han, Miguel A. Renom, Wendy E. Mackay, and Michel Beaudouin-Lafon. 2020. *Textlets: Supporting Constraints and Consistency in Text Documents*, page 1–13. Association for Computing Machinery, New York, NY, USA.

Sandra G Hart and Lowell E Staveland. 1988. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, volume 52, pages 139–183. Elsevier.

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.

Wycliffe Aber Hill. 1931. *The Plot Genie*. EE Gagnon Company.

Chieh-Yang Huang, Shih-Hong Huang, and Ting-Hao Kenneth Huang. 2020. *Heteroglossia: In-Situ Story Ideation with the Crowd*, page 1–12. Association for Computing Machinery, New York, NY, USA.

Julie S. Hui, Darren Gergle, and Elizabeth M. Gerber. 2018. *IntroAssist: A Tool to Support Writing Introductory Help Requests*, page 1–13. Association for Computing Machinery, New York, NY, USA.

Shamsi T. Iqbal, Jaime Teevan, Dan Liebling, and Anne Loomis Thompson. 2018. Multitasking with play write, a mobile microproductivity writing tool. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*, UIST '18, page 411–422, New York, NY, USA. Association for Computing Machinery.

Eric LaBouve, Erik Miller, and Foaad Khosmood. 2019. Enhancing story generation with the semantic web. In *Proceedings of the 14th International Conference on the Foundations of Digital Games*, FDG '19, New York, NY, USA. Association for Computing Machinery.

Xiaotong Liu, Anbang Xu, Zhe Liu, Yufan Guo, and Rama Akkiraju. 2019. Cognitive learning: How to become william shakespeare. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI EA '19, page 1–6, New York, NY, USA. Association for Computing Machinery.

Allan MacLean, Richard M Young, Victoria M E Bellotti, and Thomas P Moran. 1996. Questions, Options, and Criteria: Elements of Design Space Analysis. page 51.

Stephen Marche. 2021. The computers are getting better at writing. *The New Yorker*.

Hugo Nicolau, André Rodrigues, André Santos, Tiago Guerreiro, Kyle Montague, and João Guerreiro. 2019. The Design Space of Nonvisual Word Completion. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility*, pages 249–261, Pittsburgh PA USA. ACM.

Judith S. Olson, Dakuo Wang, Gary M. Olson, and Jingwen Zhang. 2017. How people write together now: Beginning the investigation with advanced undergraduates in a project course. *ACM Trans. Comput.-Hum. Interact.*, 24(1).

Hiroyuki Osone, Jun-Li Lu, and Yoichi Ochiai. 2021. *BunCho: AI Supported Story Co-Creation via Unsupervised Multitask Learning to Increase Writers' Creativity in Japanese*. Association for Computing Machinery, New York, NY, USA.

Ei Pa Pa Pe-Than, Laura Dabbish, and James Herbsleb. 2021. Open collaborative writing: Investigation of the fork-and-pull model. *Proc. ACM Hum.-Comput. Interact.*, 5(CSCW1).

Ei Pa Pa Pe-Than, Laura Dabbish, and James D. Herbsleb. 2018. Collaborative writing on github: A case study of a book project. In *Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '18, page 305–308, New York, NY, USA. Association for Computing Machinery.

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*, page 1–15. Association for Computing Machinery, New York, NY, USA.

James L. Peterson. 1980. Computer programs for detecting and correcting spelling errors. *Commun. ACM*, 23(12):676–687.

Olaf Resch and Aglika Yankova. 2019. Open knowledge interface: A digital assistant to support students in writing academic assignments. In *Proceedings of the 1st ACM SIGSOFT International Workshop on Education through Advanced Software Engineering and Artificial Intelligence*, EASEAI 2019, page 13–16, New York, NY, USA. Association for Computing Machinery.

Melissa Roemmele and Andrew S. Gordon. 2018a. Automated assistance for creative writing with an rnn language model. In *Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion*, IUI '18 Companion, New York, NY, USA. Association for Computing Machinery.

Melissa Roemmele and Andrew S. Gordon. 2018b. Automated assistance for creative writing with an rnn language model. In *Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion*, IUI '18 Companion, New York, NY, USA. Association for Computing Machinery.

K. Romer and F. Mattern. 2004. The design space of wireless sensor networks. *IEEE Wireless Communications*, 11(6):54–61.

Marlene Scardamalia and Carl Bereiter. 1987. Knowledge telling and knowledge transforming in written composition. In *Advances in applied psycholinguistics*. Cambridge University Press.

Dirk Schmidt. 2020. Grading tibetan children's literature: A test case using the nlp readability tool "dakje". *ACM Trans. Asian Low-Resour. Lang. Inf. Process.*, 19(6).

John Seabrook. 2019. The next word: Where will predictive text take us? *The New Yorker*.

Sarah Sterman, Evey Huang, Vivian Liu, and Eric Paulos. 2020. *Interacting with Literary Style through Computational Tools*, page 1–12. Association for Computing Machinery, New York, NY, USA.

Carola Strobl, Emilie Ailhaud, Kalliopi Benetos, Ann Devitt, Otto Kruse, Antje Proske, and Christian Rapp. 2019. Digital support for academic writing: A review of technologies and pedagogies. *Computers & Education*, 131:33–48.

Anne Sullivan, Mirjam Palosaari Eladhari, and Michael Cook. 2018. Tarot-based narrative generation. In *Proceedings of the 13th International Conference on the Foundations of Digital Games*, pages 1–7.

Sunny Tian, Amy X. Zhang, and David Karger. 2021. A system for interleaving discussion and summarization in online collaboration. *Proc. ACM Hum.-Comput. Interact.*, 4(CSCW3).

Xaver Tomihiro Toyozaki and Keita Watanabe. 2018. Ambientletter: Letter presentation method for discreet notification of unknown spelling when handwriting. In *The 31st Annual ACM Symposium on User Interface Software and Technology Adjunct Proceedings*, UIST '18 Adjunct, page 36–38, New York, NY, USA. Association for Computing Machinery.

Selen Türkay, Daniel Seaton, and Andrew M. Ang. 2018. Itero: A revision history analytics tool for exploring writing behavior and reflection. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI EA '18, page 1–6, New York, NY, USA. Association for Computing Machinery.

Viswanath Venkatesh and Fred D. Davis. 2000. A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science*, 46(2):186–204.

Thiemo Wambsganss, Christina Niklaus, Matthias Cetto, Matthias Söllner, Siegfried Handschuh, and Jan Marco Leimeister. 2020. *AL: An Adaptive Learning Support System for Argumentation Skills*, page 1–14. Association for Computing Machinery, New York, NY, USA.

Liuping Wang, Xiangmin Fan, Feng Tian, Lingjia Deng, Shuai Ma, Jin Huang, and Hongan Wang. 2018. Mirroru: Scaffolding emotional reflection via in-situ assessment and interactive feedback. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI EA '18, page 1–6, New York, NY, USA. Association for Computing Machinery.

Kento Watanabe, Yuichiroh Matsubayashi, Kentaro Inui, Tomoyasu Nakano, Satoru Fukayama, and Masataka Goto. 2017. Lyrisys: An interactive support system for writing lyrics based on topic transition. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces*, IUI '17, page 559–563, New York, NY, USA. Association for Computing Machinery.

Robert F. Woodbury and Andrew L. Burrow. 2006. Whither design space? *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 20(2):63–82.

Shaomei Wu, Lindsay Reynolds, Xian Li, and Francisco Guzmán. 2019. Design and evaluation of a social media writing support tool for people with dyslexia. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, page 1–14, New York, NY, USA. Association for Computing Machinery.

A Appendix

A.1 Methodology

The query we searched for searching the ACM Digital Library was:

> [[Abstract: writing] OR [Abstract: writer]] AND
>
> [[Abstract: interface] OR [Abstract: system] OR [Abstract: prototype] OR [Abstract: tool]] AND
>
> [[Abstract: assistant] OR [Abstract: support] OR [Abstract: tool]] AND
>
> [Publication Date: (01/01/2017 TO 12/31/2021)] AND
>
> [CCS 2012: Human-centered computing]

The results of the query can be found at the following url:

```
https://dl.acm.org/action/doSearch?
fillQuickSearch=false&target=advanced&
expand=dl&CCSAnd=60&AfterMonth=
1&AfterYear=2017&BeforeMonth=12&
BeforeYear=2021&AllField=Abstract%3A%
28writing+OR+writer+OR+writers%29+AND+
Abstract%3A%28interface+OR+system+OR+
prototype+OR+tool%29+AND+Abstract%3A%
28assistant+OR+support+OR+tool%29
```

Below are examples of types of papers that would or would not be included. We used these examples when determining which papers would be included.

- Some examples that would not be included: a general purpose productivity tool, where writing is an example use case; a study/analysis where the data analyzed is writing data; a study about writing-adjacent tools, like handwriting recognition; a tool that generates writing with little human interaction; a non-writing tool with a language interface; language learning tools.

- Some examples that would be included: a design fiction about a writing tool; a writing tool that has no evaluation; a writing tool that writes the first draft and then a human revises it; a study of a commercial writing tool; a tool that supports a very specific writing task; a tool that supports writing and something else (but is not a general purpose tool).

We chose this inclusion criteria subjectively, to focus on our particular interest in writing support tools and their relation to improvements in language technology. We do not intend to present this inclusion criteria as an objective definition of writing support tools. For instance, handwriting recognition may be considered a writing support tool in some contexts, but would not fit our purposes. Another small group of papers we rejected were papers that supported the collection or organization of data that would later be written about, such as a tool for quickly extracting sports-game highlights for sportswriters, and another that solicited reflections throughout the day to support memoir writing. Journalists and others may consider these writing tools, but we excluded them on the rationale that they were somewhat disconnected from the final text produced.

Table 1 shows all annotations done for the papers selected. Table 2 shows all 30 papers selected for this review, with brief descriptions and ordered by the year they were published.

There was some ambiguity in the annotations. Some tools straddled multiple parts of the writing process, or the paper didn't frame the tool in a way that clearly defined the intention of the support. Systems that provided generated text were sometimes framed as providing ideas for the writer, and these labeled as supporting 'planning', whereas others that provided generated text were framed as actually writing, and these were labeled as supporting 'translating'. However, the distinction can be subtle, and sometimes, in a user study, participants used the tool in a different way than the designers intended. Some tools had a single main feature and many small 'satellite' features, making the level of complexity unclear. Our intention with these annotations is not to provide a perfectly objective representation but rather to understand the breadth and similarities within a field of study. When an annotator was unsure about an annotation, they consulted with the rest of the team.

Some papers presented or studied more than one tool; others presented more than one evaluation for a single tool. In the case of multiple tools, we give each tool its own nickname and consider them separate entities. In the case of multiple evaluations, we consider them separate entities only when analyzing evaluation methodologies. (Multiple tools evaluated together are considered a single entity when analyzing evaluation methodologies.)

How support aligns with the cognitive process model	
part of writing process	plan / translate / review
level of constraint	1: low constraint (almost anything could be helpful) 3: medium constraint (constrained but with variety in "right" answers) 5: high constraint (support must be very specific, few "right" answers)
size of goal being support	word / sentence / paragraph / more than paragraph / writing experience
Matching creativity support tool review (Frich et al., 2019)	
complexity of tool	low: one or two features medium: multiple features, semi-complex system high: entire system or suite of tools
evaluation type	no evaluation / case study / qualitative / quantitative / mixed methods
number of participants	(numeric response)
evaluation criterion	(open response)
time spent writing with tool	(numeric response in minutes)
Quantifying type of research	
tool is exclusively about text	yes/no
tool is about collaborative writing	yes/no
tool is contribution	yes/no
technology tool uses	(open response)

Table 1: List of all annotations done for the papers. Most annotations have options, while some are open response.

Some papers studied existing commercial writing tools, and others presented novel tools developed by the researchers. The commercial writing tools studied tended to be word processors, like Microsoft Word or Google Docs. We include all of these in our analysis.

A.2 Design Space

Below are further details articulating the design space.

- Plan: Support for ideation would be included in the planning portion of the design space, as would tools that aid in structuring writing. Some brainstorming support would be lightly constrained planning, for instance during early-stage story telling, whereas other brainstorming might be highly constrained, as in when writing about historical events or in an already-constructed story world.

- Translate: We can place existing NLP tasks like automatic story generation and automatic summarization as supporting translation, where story generation tends to be only lightly constrained by a prompt and summarization is highly constrained by the document it is summarizing.

- Review: A tool that provides the writer with feedback would support reviewing, as would one that involves revising what has already been written. A lightly constrained reviewing tool might provide generic or high-level feedback like "what narrative structure are you using?" whereas a highly constrained tool might provide feedback on specific word choice, stylistic patterning, or argument coherence.

UI Design (Gonçalves and Campos, 2017): Presents a user study of four writing environments – Microsoft Word, Scrivener, OmniWriter and Ulysses. They found OmniWriter to be the most satisfying tool, and propose design guidelines for such tools, including full-screen mode for distraction-free writing.

LyriSys (Watanabe et al., 2017): Reports on a lyric generation system, which generates full song lyrics according to strain and accent constraints, and provides plenty of user control including semantic topic transitions.

Writing Together (Olson et al., 2017): Studies data traces of collaborative writing in student teams' use of Google Docs.

Liminal Triggers (Gonçalves et al., 2017): Investigates how subliminal triggering built into a writing UI may help to relieve writer's block.

GHOST (Guarneri et al., 2017): Presents a tool to support non-writers creating stories for video games. The resulting tool, GHOST, is built into Unity and aids in the creation of plot roadmaps.

Writing with RNN (Roemmele and Gordon, 2018b): Presents Creative Help, an interface that suggests new sentences in a story using an RNN language model. Study varies the degree of randomness.

MiL (Clark et al., 2018): Presents and studies creative writing support tools: a next-sentence generator for story telling, and a slogan generator for writing slogans.

AmbientLetter (Toyozaki and Watanabe, 2018): Proposes a technique to support writing activity (via autocorrection and predictive conversion) in a confidential manner with a pen-based device.

Play Write (Iqbal et al., 2018): Introduces a microproductivity tool that allows users to review and edit Word documents in small moments of spare time from their smartphone.

IntroAssist (Hui et al., 2018): Presents a tool for supporting writing introductory help requests via email by providing checklists and examples.

Itero (Türkay et al., 2018): Presents a study on how integrating writing revision analytics and visualization into writing practices can impact writing self-efficacy.

Writing on Github (Pe-Than et al., 2018): Presents the preliminary findings of a mixed-methods, case study of collaboration practices in a GitHub book project.

MirrorU (Wang et al., 2018): Presents a mobile system that supports users to reflect on and write about their daily emotional experience; provides assessment and feedback across level of detail, overall valence, and cognitive engagement.

Semantic Web (LaBouve et al., 2019): Presents a mixed initiative tool for story generation. The tool is designed to take as input a story generating grammar in addition to generic keywords and uses the semantic web to contribute real-world details.

Shakespeare (Liu et al., 2019): Presents a web application that helps with educating different writing styles through automatic style transfer (with deep learning), visual stylemotry analytics, and machine teaching (by picking out examples of a particular writing style). The authors propose a use case of this system with Shakspeare's writings.

Metaphoria (Gero and Chilton, 2019b): Presents a tool that suggests how two words might be metaphorically related.

StoryAssembler (Garbe et al., 2019): Presents StoryAssembler, an open source generative narrative system that creates dynamic choice-driven narratives, and a case study.

SMWS (Wu et al., 2019): This paper describes a tool built by the Facebook researchers to automatically 'translate' text written by people with dyslexia to non-dyslexic style writing. Having built the tool into the Facebook comment interfcae, they conduct a week long study to measure its efficacy.

Academic Writing (Resch and Yankova, 2019): Presents OKI, a chatbot tool that helps with project management, assistance in applying scientifc methods, and search in open access literature.

Style Thesaurus (Gero and Chilton, 2019a): Presents a series of automatically generated thesauruses, using word embeddings trained on custom corpuses, which are meant to reflect the stylistic preferences of the author of the corpus text.

AL (Wambsganss et al., 2020): This paper presents an NLP tool to aid student argumentative writing by providing automatic feedback on their argumentation structure.

Dakje (Schmidt, 2020): Introduces a new readability tool alongside a specific use case, and demonstrates how it can help benefit literacy in the Tibetan languages. Users have instant access to statistics on the readability of their word choices so they can make edits for easy-to-read text.

Heteroglossia (Huang et al., 2020): Presents a crowd-sourcing tool that allows writer to elicit story ideas based on a role-play strategy. The tool is developed as Google Doc add-on.

Textlets (Han et al., 2020): Introduces Textlets, interactive objects that reify text selections into persistent items, and show how Textlets can be used for selective search and replace, word count, and alternative wording.

MepsBot (Peng et al., 2020): Presents in-situ writing assistance for people commenting in online mental health communities; compares support that assesses text versus recommends text.

Literary Style (Sterman et al., 2020): Develops a model of style by training a neural net, and present novel applications including an interactive text editor with real-time style feedback.

Fork-and-Pull (Pe-Than et al., 2021): Investigates the utility of the GitHub "fork and pull" workflow for writers through a mixed-methods case study of collaborative writing. They looked at two collaborative writing cases, the first to write a mathematics textbook on homotopy type theory, and the second a set of open source public policies.

IDS System (Tian et al., 2021): This paper presents Wikum+, a website that allows you to create instances of interleaved discussion and summarization. Through a user study that compares collaborative writing with Wikum+ and with Google Docs, it was shown that Wikum+ help group writing to be more organized and to have improved coordination.

BunCho (Osone et al., 2021): Presents a tool for generating titles and synopses from keywords. Additionally, an interactive story co-creation AI system is proposed. (Japanese language)

Table 2: List of all 30 papers, ordered by the year their were published, with short description of contribution.