Can ChatGPT Do the Same?
ChatGPT and Professional Editors Compared

Daniël Janssen[1], Henri Raven[1, 2], Lisanne van Weelden[1] & Yohannes den Hertog[3]

[1]Department of Language and Communication, Utrecht University | the Netherlands

[2]Hendrikx Van der Spek Consultancy | The Netherlands

[3]HDVP Events | The Netherlands

Abstract: Since the launch of ChatGPT, the use of and debate around generative AI has grown rapidly. Professionals whose work depends on writing have expressed concern about the potential impact of such tools on their roles. But are these concerns justified? Can ChatGPT truly take on the responsibilities of a professional writer? This study investigates that question by comparing the performance of ChatGPT with that of professional editors tasked with optimizing business communication. We conducted two studies, using both qualitative and quantitative methods. In the first, three experienced editors were asked to rewrite four business letters. Their editing processes were recorded using the Microsoft Snipping Tool, and immediately afterward, we conducted retrospective interviews using stimulated recall. These interviews were transcribed and analyzed. Insights from the observations and interviews informed the design of the prompt instructions used in the second study. In the second study, we asked ChatGPT to revise the same four letters using three different prompt types. The Simple prompt instructed the model to "make this text reader-focused." The B1 prompt referred explicitly to the CEFR B1 language level, requiring ChatGPT to tailor the text for intermediate readers. Finally, the Process prompt simulated the editing steps observed in the professional editors' workflows. To evaluate outcomes, we conducted both a qualitative comparison of the revised texts and a quantitative readability analysis using LiNT, a validated tool developed for Dutch texts. Our results show that the human editors substantially improved the readability of the original letters, reducing the use of unfamiliar words, shortening complex sentences, and increasing personal engagement through pronoun use. Among the AI outputs, ChatGPT B1 achieved results most comparable to the editors, both in readability and accuracy. In contrast, ChatGPT Simple fell short in terms of clarity and introduced errors through faulty inferences. Surprisingly, ChatGPT Process also underperformed compared to ChatGPT B1 and the human editors. Only the editors' and ChatGPT B1versions were free from errors. In the discussion, we reflect on how generative AI is reshaping the concept of writing within organizations, the skills required to produce effective written communication and the impact on writing pedagogy. Rather than replacing human editors, we argue that generative AI can play a valuable role as a collaborative tool in the organizational writing process.

Keywords: AI, ChatGPT, textual characteristic, professional writing, professional editing, collaborative writing

journal of WRITING RESEARCH

Janssen, D., Raven, H., van Weelden, L., & den Hertog, Y. (2026). Can ChatGPT do the same? ChatGPT and professional editors compared. *Journal of Writing Research*, *17(3),* 387-417. DOI: https://doi.org/10.17239/jowr-2026.17.03.02

Contact: Daniel Janssen, Department of Language and Communication, Utrecht University | the Netherlands - d.m.l.janssen@uu.nl

1. Introduction

The rise of artificial intelligence (AI), particularly tools like ChatGPT, is profoundly affecting professions centered on text. In the Netherlands, professional translators recently raised concerns in national media, citing fewer assignments and declining rates. “Rates are just plummeting, and AI is the big culprit,” says translator Marianne van Amersfoort-Gerritsen. “Clients increasingly use AI themselves, which not only reduces demand but shifts the focus from translation to editing AI-generated texts. That pays about half as much, yet is just as labor-intensive” (NOS News, 9/7/2024).

The spread of generative AI is causing anxiety beyond translation, affecting other professional (text)writers. Companies now use AI for business correspondence, chats, marketing, and even creative content, putting growing pressure on copywriters. Their craftsmanship and originality are increasingly undervalued in a market that prioritizes speed and cost. Henry Williams' article headline in *The Guardian* (January 24, 2023) captures this mood: *“I'm a copywriter. I'm pretty sure artificial intelligence is going to take my job.”* Not all voices are as pessimistic. Teddy Stevens, writing in *Medium* (March 13, 2023), argues that while AI can assemble a story, it lacks the depth, emotion, and human experience needed to create something like a good novel.

Still, new technologies are changing how we create texts—and how we understand writing itself. This is not unprecedented. Luuk van Waes (1991) demonstrated how word processors influence writing processes (see also Heilmann, 2023). Mariëlle Leijten (2007) studied the effects of speech recognition, and with John Hayes and Karen Shriver (2014), explored how internet access reshapes writing. Today, generative AI (cf. Benites et al., 2023) introduces fresh challenges. For writing researchers, this technological shift opens new avenues for study and innovation. One focus is how tools like ChatGPT affect writing processes, as in Noy and Zhang’s (2023) experiment testing whether ChatGPT enhances work efficiency and text quality among professionals. (We examine their findings in more detail later.)

Our research aligns with Noy and Zhang’s focus on professional writing—but differs in its emphasis on revision processes. We studied three professional editors tasked with rewriting four organizational texts to improve clarity for their target audience. Their editing sessions were recorded and analyzed in consultation with them. Next, ChatGPT rewrote the same texts using three different prompts, one of which aimed to replicate the editors’ process. We then conducted both quantitative and qualitative comparisons of all versions.

This paper is structured as follows: Section 2 situates our experiment within broader research on professional (re)writing, framing ChatGPT as a generative AI tool for such tasks. Section 3 outlines our experimental design and methodology. Section 4 offers a qualitative account of the editors’ and ChatGPT’s performances. Section 5 presents a quantitative comparison. Finally, Section 6 addresses study limitations, key conclusions, and implications for research and practice.

2. Context of the study

2.1 Professional writing

Our research focuses on professional editors who work on business texts—an area that has seen relatively little attention in writing research in recent decades. Most studies center on educational contexts and improving students' writing skills. Yet, interest in business writing is not new. In the 1980s, scholars like Jack Selzer (1983, 1984) critiqued Flower and Hayes' (1981) model for its narrow applicability, noting significant differences from professional writing (cf. Barabas, 1990). Selzer observed engineers in organizations drafting business reports in highly linear, non-recursive processes—unlike those described by Flower and Hayes (see also Hayes, 2012).

From the late 1980s to early 2000s, writing research also explored work-related writing processes (e.g., Odell & Goswami, 1985; Kogen, 1989). The field of professional writing emerged in part because organizations began to recognize the value of effective writing for operational success (Barabas, 1990; Duin, 1991; Janssen & Neutelings, 2001).

A key insight from this research is that organizational writing is a social activity shaped by institutional structures. Business writing typically involves multiple stakeholders and unfolds through cycles of production, commentary, and revision (Ede, 1990; Lay & Karis, 1991). A professional writer's effectiveness depends not only on language skills but also on managing interactions, structuring writing processes, navigating power dynamics, and adapting to organizational cultures (Janssen, 2001; Van der Mast & Janssen, 2001).

These distinctions are evident in research on policy writing (Janssen & Van der Mast, 1991; Van der Mast & Janssen, 1991). Policy texts emerge from negotiation among diverse interest groups and aim to reflect a consensus. To accommodate multiple interpretations, they are often deliberately vague or multi-layered. In such cases, acceptability may take precedence over readability.

Another key difference between business and educational writing lies in ownership and intent. Organizational texts are collective outputs that serve specific functional purposes and belong to the organization, unlike student texts, which are personal and expressive. Business writing is meant to facilitate operations and often projects a corporate identity (Pander Maat & Steehouder, 1992; Janssen & Schilperoord, 1992; Barabas, 1990). For more on corporate identity in texts, see Van Riel & Fombrun (2007), Anzidei (2002), and Horning (2006).

In sum, writing for or on behalf of an organization is fundamentally different from producing an academic paper or a personal article, such as one for a general-interest magazine (Flower & Hayes, 1981). This insight has shaped new writing pedagogies and models, including the role of collaboration and peer feedback in writing instruction (Lundstrom & Baker, 2009; Van Steendam, 2016) and Graham's (2018) writers-within-community theory. Hayes' (2012) updated model reflects this evolution, incorporating "collaboration and critics" and "technology" into the writing environment. Leijten et al. (2014) further developed this model through a case study that connects traditional writing theories to the digital age, where writers increasingly rely on diverse digital sources in their work.

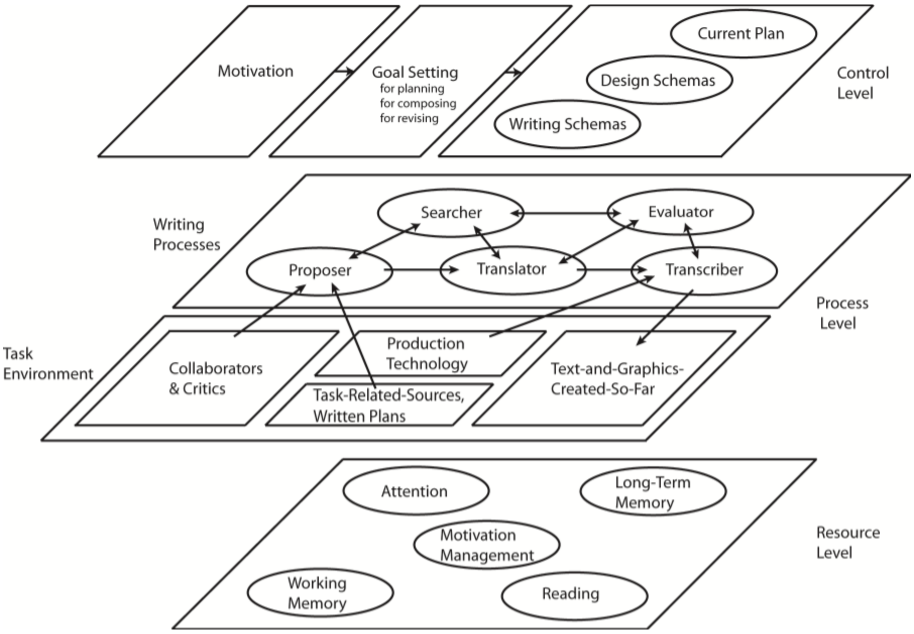

Figure 1: Model of composing (from Hayes 2012) elaborated to encompass activities of skilled professional communicators (from: Leijten et al. 2014)

Figure 1 (Leijten et al., 2014) presents a model of professional writing that emphasizes the integration of multiple digital sources. In this approach, professionals collaboratively and dynamically construct documents by combining their own content with external text, graphics, and digital materials. Instead of relying solely on memory or a single source, writers navigate various tools and resources—such as previous documents, templates, emails, websites, and other assets—to create coherent outputs. This process involves extensive searching, paraphrasing, and adapting digital content to meet specific objectives.
Professional writing, in this model, is both cognitively demanding and socially interactive. Writers must address rhetorical goals while coordinating with colleagues and external contributors. Internet sources are particularly central, enabling professionals to search online, collaborate via platforms like Google Docs, consult company wikis for historical information, and verify content by tracing its origins. Visual elements, including charts and diagrams, are incorporated and adapted to enhance clarity and impact.

Leijten et al. demonstrate how Inputlog, a keystroke-logging tool, captures interactions between writers and their digital environments. It reveals patterns of content searching, reuse, and the iterative nature of writing, where professionals draft, refine, and revisit multiple sources. Their case study illustrates how writers fluidly switch between resources, integrate varied content, and refine proposals—highlighting the complexity and effectiveness of digital-era workflows.

Within this model, ChatGPT—or any large language model—can support the collaborative, step-by-step parts of professional writing. It can draft, rephrase, and adapt text while combining input from different contributors. Because these models are fine-tuned to follow instructions and trained with human feedback (RLHF), they can follow style and tone guidelines which enables audience-aware rewriting (Ouyang et al., 2022; Longpre et al., 2023). By offering suggestions and refining drafts, ChatGPT may assist cognitive writing processes such as brainstorming, rhetorical tailoring, and audience-specific customization (Flower & Hayes, 1981; Kellogg, 2008; Lee et al., 2022; Yuan et al., 2022.

In professional workflows, ChatGPT might act as more than a tool—potentially functioning as a virtual collaborator. It can respond immediately to prompts, adapt its output to evolving needs, and contribute to iterative revisions, much like a human partner (Lee et al., 2022; Yuan et al., 2022). Its ability to generate ideas, suggest structural improvements, and contextualize information can boost productivity for individuals and teams; controlled studies find sizable gains on professional writing tasks and in customer-support writing (Noy & Zhang, 2023; Brynjolfsson et al., 2023). Additionally, by simulating virtual searches and offering targeted recommendations (via retrieval-augmented generation or tool-use agents), ChatGPT may take on the role of a knowledgeable colleague, guiding writers toward clearer communication and better-informed decisions (Lewis et al., 2020; Yao et al., 2022.

ChatGPT may also assist in structuring documents by suggesting formats, outlines, and layouts aligned with rhetorical goals and audience expectations; prototype systems show measurable benefits for planning, outlining, and rapid draft prototyping (Zhang et al., 2023; Lee et al., 2024). If integrated into platforms like Google Docs, it could enhance collaborative environments by complementing existing tools with its language-generation capabilities (Google, 2023).

2.2 Professional editing

This contribution focuses on a specific aspect of organizational writing: professional editing. As noted earlier, writing tasks in organizations are often distributed according to employees' expertise, roles, and positions. Sometimes, different departments contribute text proposals that are later integrated, as in annual reports (Cross, 1990). Writing tasks—ranging from information gathering to drafting, editing, and revising—are commonly divided among departments and staff (Barabas, 1990).

To manage these processes, organizations sometimes hire external experts. This occurs when they have content expertise but lack communication skills, when internal conflicts arise, or when time or workload constraints prevent internal handling. In such cases, professional editors or agencies are brought in to enhance readability and accessibility (Bissaillon, 2007).

Hiring external editors offers several advantages. They bring an objective perspective—vital when multiple contributors make it hard to identify the core message—or when subject matter experts accurately present content but struggle to make it accessible. External editors also offer specialized skills to simplify complex technical or legal information or tailor a document's style to organizational standards (e.g., house style). Additionally, they offer valuable outside perspectives during revision—an improvement over self-revision, which many internal writers rely on (Willey & Tanimoto, 2012). Typically, writers follow the revision steps described by Flower et al. (1987): detection, diagnosis, and revision. At the final stage, authors may revise, postpone, or ignore identified issues. While professional editors also detect and diagnose issues, they aren't required to solve them directly. Instead, they communicate these findings to the author, organization, or client. Willey and Tanimoto identify two strategies in such cases: editors can either return the issue to the original author(s), shifting responsibility, or request input to collaboratively resolve it (cf. Burrough-Boenisch, 2006). By clearly identifying issues and discussing them with the organization, editors help harness collective expertise to optimize texts.

2.3 ChatGPT as a writing or editing tool

As previously mentioned, professionals in the writing and editing industry are concerned that generative AI like ChatGPT may replace their expertise. ChatGPT produces "natural language" responses to user input, enabling it to rewrite text in ways similar to professional writers and editors.

ChatGPT is built on a Large Language Model trained on vast text datasets to recognize, understand, and generate language (Wikipedia, n.d.; Markowitz, 2021). It interprets plain-language prompts using syntactic, pragmatic, and semantic analysis, drawing on both direct input and contextual cues.

In generating text, ChatGPT predicts words sequentially based on learned patterns, producing fluent and grammatically correct responses. Its training enables it to adapt to different styles and contexts—for example, using an informal tone for advertising or a formal style for scientific writing.

However, ChatGPT has notable limitations. One issue is the generation of false information—referred to as "hallucinations" by Alkaissi and McFarlane (2023). When we asked ChatGPT for quotes from copywriters concerned about AI, it returned fabricated quotes and sources. ChatGPT also provides inconsistent responses to identical prompts, which may be problematic in professional settings. Its non-deterministic nature makes uniform output unattainable. Moreover, as it is predominantly trained on English texts, its performance in other languages (e.g., Dutch) is limited—less than 1% of its training data is estimated to be Dutch.

By late 2023 - early 2024 when we conducted this study, ChatGPT was the dominant, most widely accessible writing assistant: it was estimated to have reached aproximately 100 million monthly users by January 2023 and, beyond the web UI, offered official access routes (API; iOS app) that made it practical for real workflows (Reuters, 2023; OpenAI, 2023a; OpenAI, 2023b). Other models were available but with caveats. Claude 2 (Anthropic) launched in July 2023 with very long context windows and careful, tone-preserving edits, but early availability was limited to the U.S. and U.K. via web and a paid API in limited access (Wiggers, 2023; Anthropic, 2023; Reuters, 2023). Mistral focused on open-weight developer-centric releases—Mistral 7B (Sept 27, 2023) and the larger Mixtral 8×7B (Dec 11, 2023), which appealed for on-premise or privacy-sensitive use but were less turnkey for non-technical teams at that time (Mistral AI, 2023a; Mistral AI, 2023b). Google's Bard expanded to 40+ languages and more countries in July 2023, lowering barriers for everyday drafting inside Google's ecosystem, though its role in professional writing teams was still emerging (Google, 2023). In this landscape, selecting ChatGPT for our study let us evaluate a widely adopted, easily accessible baseline that professionals could realistically have used.

To date, no experimental research has examined ChatGPT's role in enhancing business texts edited by professionals, as proposed in this study. However, recent experiments have compared ChatGPT-generated texts with those written by students (Herbold et al., 2023) and professionals (Noy & Zhang, 2023).

In Herbold et al.'s (2023) study, 90 students wrote argumentative essays, which were compared to those generated by ChatGPT-3 and ChatGPT-4. Teachers assessed text quality and analyzed lexical and grammatical features. ChatGPT's essays had more grammatical and stylistic errors but showed greater syntactic complexity, formal tone, and lexical diversity. AI-generated texts were also more structured, though students used more modal verbs ("should," "can") and epistemic expressions ("I think"), creating a more personal, nuanced tone. While ChatGPT used standard phrases (e.g., "In conclusion"), student essays were freer and less formulaic. Linguistic analysis confirmed these patterns, with ChatGPT showing more complexity and nominalizations, and students using more discourse markers ("on the other side," "moreover").

Noy and Zhang's (2023) study involved 435 professionals—marketers, consultants, data analysts, and grant writers—who completed two tasks: a press release and an annual report. For Task 1, participants were split into a control group and a ChatGPT-3.5 group. For Task 2, all participants could use ChatGPT. Researchers measured time and text quality (via jury ratings). Those who used ChatGPT in Task 1 were more likely to use it again, indicating a positive experience. ChatGPT users completed tasks 40% faster, and their texts were rated 18% higher in quality. Less experienced writers benefited the most, suggesting ChatGPT helps reduce skill-based disparities—a so-called Matthew effect.

Together, the studies by Herbold et al. (2023) and Noy & Zhang (2023) show that ChatGPT can enhance writing processes, text characteristics, and overall quality. The main drawback is its lack of personality.

In this study, we aim to extend those findings by (1) comparing ChatGPT's performance with that of professional editors, and (2) focusing specifically on the revision or editing processes.

3. Methodology

In this section, we discuss the design and implementation of the study. We describe who our subjects were and what kind of rewriting tasks we gave them to do. We will also show how we mapped the rewriting processes and how we compared the results, the rewritten texts, with each other. Finally, we justify what rewriting prompts we gave ChatGPT and why.

3.1 Subjects and tasks

In this study, we first asked 3 professional editors (Ann, Derek, and Hank) to rewrite 4 letters. The editors are employed by the same agency in the Netherlands, and all have extensive experience with this task (20+ years) and with business texts. When selecting the texts, we considered the length of the text (it had to be limited) and the variation in the subject, the sender and the recipient (they had to be slightly different). The limitation was necessary to keep the task feasible for the editors. Varying is also essential to ensure external validity.

To keep the rewriting tasks manageable, we selected letters/emails that could be rewritten in a limited amount of time. The longest text consisted of 29 sentences and the shortest of 13 sentences. The texts came from 4 different organizations, dealt with different topics, and had different audiences (see figure 2). Thus, we could expect at least some variation in revision processes.

Allianz	Allianz is an international insurance and financial services company. The letter from Allianz is an informational letter asking the reader to return information to Allianz. The reader is a customer insured with Allianz. The letter informs the reader of a prior request to discontinue the reader's life insurance policy. This request was made by the reader's spouse. In order to process this request, Allianz asks the reader to send some documents.
Liander	Liander is a grid operator. The letter from Liander is an informational letter and the reader is a resident of a street where Liander provides grid management. The reader is informed about work that will take place in the street. In the letter Liander explains what work will take place and what it means for the reader.
Keerpunt	Keerpunt is a specialist in the field of reintegration and occupational health and safety services. Keerpunt's letter is an informational letter instructing the reader to apply for benefits. The reader is an employer who has registered a pregnant employee with the UWV. The letter explains that the employer must apply for WAZO benefits for the employee. That letter explains how to do this, what happens if the employee gets sick before the leave period, and when the employee is entitled to which benefit (WAZO benefit or Sickness Benefits Act benefit).
Zorg en Zekerheid	Zorg en Zekerheid is a regional health insurer based in Leiden. The letter from Zorg en Zekerheid is an informational letter with a request to send information. The letter is addressed to the practitioner of a Zorg en Zekerheid customer and was written in response to a reimbursement request from the practitioner. This request lacks information that Zorg en Zekerheid needs to process the request. Therefore, Zorg en Zekerheid requests the missing information through this letter.

Figure 2: Material

Our editors were given the task: "Make this letter/email a good letter/email”. We chose this broad description because clients never give professional editors more specific instructions than this. They trust that the editor will know better than they themselves what the requirements are for a good letter or e-mail.

3.2 Recording and analysis

The editors completed the task independently in their own offices, working in Microsoft Word. Their writing processes were recorded using the Microsoft Clipping Program, a built-in screen recorder available on Windows 10 and later. This tool captures all on-screen activity without interfering with the writer's workflow. Immediately after the task, we conducted retrospective interviews using a stimulated recall method (Calderhead, 1981). The interviews were transcribed with Amberscript and summarized by ChatGPT to identify patterns and differences between the three editors. We verified these summaries against our own notes and found them to be consistent.

We analyzed the rewritten texts both qualitatively and quantitatively. For the qualitative analysis, we examined the revisions made by both the human editors and ChatGPT. Due to space constraints, we focus here on a single letter (Keerpunt), one editor (Ann), and one prompt condition (ChatGPT Simple). For the quantitative analysis, we used LiNT, an online tool developed at Dutch universities to assess the readability of Dutch texts through automatic syntactic, pragmatic, and semantic analysis (see Benites et al., 2023). LiNT has been extensively validated, especially in Kleijn (2018), and provides scores for 'word difficulty,' 'sentence difficulty,' and 'personality.'

Word difficulty is determined by factors such as word familiarity, abstract nouns, and the number of new word occurrences. Sentence difficulty includes sentence length, phrase complexity, enumerations, and use of subordinate clauses. Personality is assessed based on references to people, personal pronouns, direct address, and active voice. For further details on these variables and scoring, see Kleijn (2018) and the LiNT website (https://lint.hum.uu.nl/uitleg/tekst-kenmerken).

3.3 Prompting ChatGPT

In the second part of our study, we asked ChatGPT to rewrite the same four texts using three different prompts: a *simple* prompt, a *B1* prompt, and a *process-based* prompt. We chose to vary the prompts because ChatGPT's output is highly sensitive to the specificity and quality of the input it receives. Each prompt reflected a different level of complexity, allowing us to explore how varying instructions influence the model's performance.

The *simple* prompt—"Make this text reader-friendly"—is similar in spirit to the general instruction given to the human editors, who also aimed to improve readability rather than merely correct errors. The *B1* prompt asked ChatGPT to "Rewrite this text to B1 level," referring to the language proficiency level defined by the Common European Framework of Reference for Languages (CEFR). Although the instruction may appear straightforward, B1-level writing requires implicit knowledge of linguistic simplification, including the use of familiar vocabulary and clear sentence structures—something not easily achieved without deeper understanding.

Finally, we developed a *process* prompt modeled after the step-by-step rewriting practices observed in the professional editors. This was our most detailed and specialized instruction.

ChatGPT was asked to perform the revision in eight distinct steps, mirroring the expert workflow of our editors and integrating more advanced knowledge of editing strategies:

1. Scan the letter and identify the problems that make it not a good letter.
2. Identify the purpose, topic, audience, and central question.
3. Gather all relevant information from the letter
4. Organize the relevant information pyramidally
5. Rewrite the letter, making sure you write in an attractive, understandable, and correct manner.
6. Pay special attention to the following five style problems when formulating: auxiliary verbs, nominal style, passive voice, split or brace constructions/embedded clauses, and prepositional phrases.
7. Determine the CERF (Common European Framework of Reference for Languages) level of the letter, and when rewriting, make sure it is B1 level.
8. Perform a final check to verify that the text is now optimized.

4. Results 1: qualitative comparison of human and machine rewrites

To explore differences in editorial quality between a professional human editor and ChatGPT, we focus on a single letter rewritten by editor Ann and by ChatGPT using the basic prompt ChatGPT Simple. As mentioned, we limit our comparison to one of the four original letters—Keerpunt—as it was the longest and most representative of the types of rewriting tasks professional editors typically face. Ann's rewriting is representative of the work produced by the other professional editors involved. We selected ChatGPT Simple because it reflects the kind of minimal prompting likely used by lay users. The outcomes of Ann and ChatGPT are discussed in contrast with results from the other rewriting 'conditions.' Figure 3 shows the original Keerpunt letter and a literal translation.

	Original	Original (translated)
1.	Betreft: Zwangerschapsmelding	Subject: Notification of pregnancy
2.	Geachte heer/mevrouw,	Esteemed Sir/Madam,
3.	Naar aanleiding van de zwangerschapsmelding begrijp ik dat uw werknemer XX met verlof zal gaan. Via deze weg wil ik u erop wijzen dat u een WAZO-uitkering dient aan te vragen bij het UWV.	Following the pregnancy notification, I understand that your employee XX will be going on leave. This way, I would like to remind you that you need to apply for a WAZO benefit through the UWV.
4.	*WAZO uitkering aanvragen* Met deze link kunt u de zwangerschap doorgeven aan het UWV en de WAZO uitkering aanvragen. Dit kan maximaal 4 en minimaal 2 weken voor de ingangsdatum	*Applying for WAZO benefit* With this link, you can report the pregnancy to the UWV and apply for the WAZO benefit. This can be done a maximum of 4 weeks and a minimum of 2 weeks before the start date of

van het verlof. Uw zwangere werknemer heeft recht op minimaal 16 weken zwangerschapsverlof, waarbij het verlof na de bevalling nog altijd nog minimaal 10 weken duurt. Tijdens dit verlof betaalt UWV een zwangerschapsuitkering.

the leave. Your pregnant employee is entitled to a minimum of 16 weeks of maternity leave, with the leave still lasting at least 10 weeks after childbirth. During this leave, the UWV will pay a maternity benefit.

5. Wij raden een werkgever aan om concrete afspraken te maken over het moment van hervatten na het verlof en over het aantal contracturen.

We recommend that an employer makes concrete arrangements regarding the timing of the return after leave and the number of contract hours.

6. *Ziektewetuitkering aanvragen bij (gedeeltelijk) uitval vóór of na verlof*
Het kan voorkomen dat uw werknemer eerder uitvalt a.g.v. zwangerschap gerelateerde klachten of na verlof niet (volledig) kan hervatten. Zij heeft dan mogelijk recht op een Ziektewetuitkering. Als werkgever vraagt u voor haar de Ziektewetuitkering aan. U kunt dit digitaal aanvragen. Tevens dient u hier een melding van te maken in Verzuimsignaal.

Applying for Sickness benefit in case of (partial) absence before or after leave
It may happen that your employee becomes unable to work earlier, due to pregnancy-related complaints or cannot fully return after leave. In this case, she may be entitled to a Sickness benefit. As the employer, you will need to apply for the Sickness benefit on her behalf. This can be done digitally. You should also make a notification in Verzuimsignaal.

7. Is uw werknemer weer (gedeeltelijk) aan het werk of weer volledig ziek, meldt dit dan ook bij het UWV.

If your employee is (partially) back at work or is fully sick again, please report this to the UWV as well.

8. Let op! Afhankelijk van de periode wanneer uw medewerkster ziek geworden is, wordt er bepaald op welke uitkering zij recht heeft. Is zij 6-4 weken vóór de vermoedelijke bevallingdatum nog ziek maar is haar zwangerschapsverlof nog niet begonnen? Dan krijgt zij een Ziektewetuitkering en deze ziektedagen worden van het zwangerschapsverlof van minimaal 16 weken afgetrokken. Op de website van het UWV vindt u hier meer informatie over.

Note! Depending on the period when your employee became sick, it will be determined what benefit she is entitled to. If she is still sick 6-4 weeks before the expected delivery date but has not yet started her maternity leave, she will receive a Sickness benefit, and these sick days will be deducted from the minimum 16 weeks of maternity leave. You can find more information about this on the UWV website.

9. *Heeft u nog vragen?*
Heeft u nog vragen betreffende de melding bij het UWV? Neem dan contact op met het UWV (0900-9295). Heeft u vragen betreffende de arbodienstverlening of Verzuimsignaal? Neem dan gerust contact met mij op. Ik ben bereikbaar op 035-

Do you have any questions?
If you have any questions regarding the notification to the UWV, please contact the UWV (0900-9295). If you have questions about occupational health services or Verzuimsignaal, please feel free to contact me.

	6253147. U kunt mij ook een e-mail sturen: s.hekker@keerpunt.nl.	I can be reached at 035-6253147. You can also send me an email at s.hekker@keerpunt.nl.
10.	Met vriendelijke groet, Keerpunt b.v. NAAM	Yours sincerely, Keerpunt B.V. NAME

Figure 3: Original Keerpunt letter and translation

4.1 Human editors

We begin by examining the revision produced by editor Ann (see Figure 4.). The original Dutch letter is presented below, followed by an English translation. Since our analysis focuses on the Dutch rewritings, the translation is included solely to support non-Dutch-speaking readers in understanding the source text. It aims to stay as close as possible to the original in both content and tone, without introducing interpretative changes.

	Ann's letter	Ann's letter (translated)
1.	**Betreft**: Zwangerschapsmelding	**Subject**: Pregnancy notification
2.	Geachte heer/mevrouw XX,	Dear Mr./Ms. XX,
3.	Hartelijk dank voor de zwangerschapsmelding van uw werknemer XX. Ik begrijp dat deze werknemer met verlof gaat. Het is daarom verplicht een WAZO-uitkering aan te vragen bij het UWV. In deze e-mail leest u daar meer over.	Thank you very much for your employee XX's pregnancy notification. I understand that this employee is going on leave. It is therefore mandatory to apply for WAZO benefits at the UWV. You can read more about this in this e-mail.
4.	**Wat is een WAZO-uitkering?** Het UWV betaalt aan u een WAZO-uitkering gedurende het zwangerschapsverlof van uw medewerker. Uw zwangere werknemer heeft recht op minimaal 16 weken zwangerschapsverlof: 6 weken voor en 10 weken na de bevalling. Als uw werknemer dus later dan de uitgerekende datum bevalt, wordt het verlof langer. Na de bevalling heeft zij namelijk sowieso recht op 10 weken verlof.	**What is a WAZO benefit?** The UWV will pay WAZO benefits to you during your employee's maternity leave. Your pregnant employee is entitled to at least 16 weeks of maternity leave: 6 weeks before and 10 weeks after giving birth. Therefore, if your employee gives birth later than the due date, the leave will be longer. In fact, after giving birth, she is entitled to 10 weeks of leave anyway.

5.	**Hoe vraagt u de WAZO-uitkering aan?** Met deze link kunt u de zwangerschap doorgeven aan het UWV en de WAZO-uitkering aanvragen. Dit kan maximaal 4 en minimaal 2 weken voor de ingangsdatum van het verlof.	**How do you apply for WAZO benefits?** You can use this link to report the pregnancy to the UWV and apply for WAZO benefits. This can be done up to 4 and at least 2 weeks before the effective date of the leave.
6.	**Ziektewetuitkering aanvragen bij (gedeeltelijk) uitval vóór of na verlof** Het kan voorkomen dat uw werknemer eerder uitvalt of dat zij na het verlof haar werk niet (volledig) kan hervatten. Zij heeft dan mogelijk recht op een Ziektewetuitkering. Als werkgever vraagt u voor haar de Ziektewetuitkering aan. U kunt dit digitaal aanvragen. U maakt hier dan een melding van in Verzuimsignaal.	**Apply for sickness benefit in case of (partial) failure before or after leave** It may happen that your employee drops out early or that she is unable to resume her work (in full) after the leave. She may then be entitled to benefit under the Sickness Benefits Act. As the employer, you apply for sickness benefit on her behalf. You can apply for this digitally. You then make a notification of this in Verzuimsignaal.
7.	**Is uw werknemer weer (gedeeltelijk) aan het werk of weer volledig ziek?** In dat geval meldt u dat ook bij het UWV.	**Is your employee back at work (partially) or fully sick again?** In that case, you also report that to the UWV.
8.	**Op welke uitkering heeft uw werknemer recht?** Afhankelijk van de periode wanneer uw werknemer ziek geworden is, bepalen we op welke uitkering zij recht heeft. Is zij 6 tot 4 weken vóór de vermoedelijke bevallingdatum nog ziek maar is haar zwangerschapsverlof nog niet begonnen? Dan krijgt zij een ziektewetuitkering; deze ziektedagen worden dan van het zwangerschapsverlof van minimaal 16 weken afgetrokken. Op de website van	**What benefit is your employee entitled to?** Depending on the period when your employee became sick, we determine what benefit she is entitled to. Is she still sick 6 to 4 weeks before the expected delivery date but her maternity leave has not yet started? Then she receives sickness benefit; these sick days are then deducted from the maternity leave of at least 16 weeks. You can find more information about this on the UWV website.

	het UWV vindt u hier meer informatie over.	
9.	**Heeft u vragen over deze e-mail?**	**Do you have any questions about this e-mail?**
	Neemt u dan contact op met het UWV (0900-9295). Voor vragen over de arbodienstverlening of Verzuimsignaal kunt u rechtstreeks contact met mij opnemen. Ik ben bereikbaar op 035-6253147. U kunt mij ook een e-mail sturen: s.hekker@keerpunt.nl. Ik help u graag verder.	If so, please contact the UWV (0900-9295). For questions about health and safety services or Verzuimsignaal, please contact me directly. I can be reached at 035-6253147. You can also send me an e-mail: s.hekker@keerpunt.nl. I will be happy to help you.
10.	Met vriendelijke groet, Keerpunt b.v. NAAM	Best regards, Keerpunt b.v. NAME

Figure 4: Ann's rewrite

Ann made a range of structural and stylistic improvements to the original Dutch version of the *Keerpunt* letter. While the subject line remained unchanged, she personalized the salutation slightly by inserting "XX" to reflect anonymized personalization. The opening paragraph (3) was notably restructured: instead of a single, somewhat formal sentence, Ann split the message into three concise and friendly sentences. She added a brief thank-you and removed the phrase "via deze weg wil ik u erop wijzen" ("this way I would like to remind you") which was unnecessarily formal and indirect. Importantly, this revised opening now more clearly communicates the purpose of the letter, making it immediately apparent to the reader what action is required and why the message is relevant. To improve readability, Ann added informative section headers throughout the letter. For example, she introduced the heading "Wat is een WAZO-uitkering?" ("What is a WAZO benefit?"), followed by a clearer and more accessible explanation of the benefit and its conditions. She added a useful nuance about how the leave period is extended if the employee gives birth later than expected.

The section explaining how to apply for the WAZO benefit was refined, now preceded by the heading "Hoe vraagt u de WAZO-uitkering aan?" ("How do you apply for WAZO benefits?"). This makes the structure of the letter more navigable for readers.

Ann also simplified and clarified the section on sickness benefits (4). She replaced formal language (e.g. abbreviations like "a.g.v.") with more straightforward wording and added a clear heading: "Ziektewetuitkering aanvragen bij (gedeeltelijk) uitval vóór of na verlof" ("Apply for sickness benefit in case of (partial) failure before or after leave"). Sentences in this section (6) were shortened and rephrased for clarity and flow. The passage on reporting an employee's return to work or renewed absence (7) was edited to be more concise and direct, with the addition of a helpful subheading. Likewise, the explanation (8) about the employee's entitlement to sickness benefits depending on the timing of illness was improved: the informal attention marker "Let op!" ("Note") was removed, the language simplified, and the logic clarified. In the closing section (9), Ann enhanced the customer-oriented tone. Instead of separating questions based on their topic (UWV vs. occupational health), she presented a more unified, approachable contact paragraph, ending with "Ik help u graag" ("I will be happy to help you").

Overall, Ann improved the structure, tone, and clarity of the original letter significantly. Her revision can be characterized as more user-friendly, professional, and accessible, while preserving the original content.

Ann's rewrite aligns with those of her fellow professional editors, Hank and Derek (see Appendix 1). All three improve upon the original letter by simplifying the language, breaking up dense passages, and using subheadings to enhance readability. They retain the original content and sequencing while outlining employer responsibilities such as applying for WAZO benefits and reporting (partial) sick leave. Each version clarifies that maternity leave may extend beyond 16 weeks and that the employer plays a central role in initiating benefit applications.

The main differences lie in tone and phrasing. Ann adopts a warmer, more client-friendly approach, opening with a thank-you and using supportive, accessible language. For instance, she begins the letter with *"Hartelijk dank voor de zwangerschapsmelding van uw werknemer XX"* ("Thank you very much for your employee XX's pregnancy notification"), which sets a welcoming tone. In contrast, Derek opens more formally with *"Wij hebben een zwangerschapsmelding ontvangen voor uw werknemer XX"* ("We have received a pregnancy notification for your employee XX"), signaling a more institutional voice. Similarly, while Ann avoids bureaucratic expressions, Hank retains more formal phrasing, such as *"Let op! Afhankelijk van de periode wanneer uw medewerkster ziek geworden is..."* ("Note! Depending on the period when your employee became sick…"), which emphasizes regulatory compliance. Derek takes a comparable approach, using more technical instructions like *"Leg de melding ook vast in Verzuimsignaal"* ("Record the notification in Verzuimsignaal as well"), whereas Ann rephrases such guidance in more reader-friendly terms.

Despite these variations, all three rewritings result in well-structured, informative, and user-oriented letters. Their differences are primarily stylistic, while their editorial decisions consistently reflect a shared focus on clarity, usability, and audience relevance.

Our retrospective interviews with editors revealed several similarities in their approach, along with some minor individual differences. All three begin by reading the original letter in full to get a sense of its overall structure, tone, and content before making any detailed changes. This initial scan serves to establish a global understanding of the message. As Derek explains, *"I read it through first. In its entirety […]. I just want to get a quick feel for the content of the letter."* Ann describes a similar process: *"I thought I'd read it through once anyway."*

A shared focus among the editors is improving functionality, tone and clarity. Each aims to replace formal language with simpler, more accessible phrasing. They also note the importance of adapting a tone to match contemporary communication styles. Throughout, all three editors keep the target audience in mind and aim to revise the text to meet B1-level readability standards. Attention to structure is another point of convergence. The editors restructure the letters to foreground essential information and improve flow. Ann explicitly refers to this as a *"pyramid structure"* approach: *"So I'm looking for the most important message. I put that at the top."* Hank and Derek also adjust the order and hierarchy of the content, omitting redundant details to sharpen the core message.

One notable difference concerns the extent to which the editors reflect on the broader communicative context. Hank, for example, highlights that in practice he would often question the necessity of the letter itself: *"In a real situation, I would say to Keerpunt, 'Should you send this letter?'"* He emphasizes that editorial work often occurs in collaboration with the client, involving not just textual revision but also joint decision-making about the appropriateness and strategic purpose of the communication. In this way, editors act not only as language specialists but also as communication advisors.

4.2 ChatGPT

Having outlined the shared practices and editorial choices of the professional editors, we now turn to the output generated by *ChatGPT Simple*—the most basic prompt condition used in this study. This comparison allows us to examine how a minimally instructed language model handles the same rewriting task and to what extent its approach and output align with, or diverge from, those of human professionals. By contrasting *ChatGPT Simple* with the revisions by Ann, Hank, and Derek, we can better understand the strengths and limitations of the model in relation to real-world editorial expertise.

Figure 8 presents the rewrite produced by ChatGPT under the *Simple* prompt condition.The first notable change is in segment 1, where the subject line has been expanded to better reflect the content of the letter. Segment 3, however, contains some problematic revisions. The letter opens with the line, *"Gefeliciteerd met de aanstaande uitbreiding van uw team"* ("Congratulations on the upcoming expansion of your team") which is inappropriate: the employee is going on maternity leave, and the baby is not joining the team as a new employee. This represents an inferential error likely caused by ChatGPT drawing from reference texts in which "pregnancy" and "congratulations" frequently co-occur, but in this context, the phrasing is misleading. Additionally, the segment refers to "*stappen om een WAZO uitkering aan te vragen*" ("steps to apply for WAZO benefits") even though the application process consists of only a single step. The remainder of the letter covers related but distinct actions the employer may need to take. The closing phrase, "*zorgen dat alles soepel verloopt*" ("to ensure everything goes smoothly") is vague, for instance raising questions about what "everything" refers to.

Segment 4 contains a complex brace construction that hampers readability. The sentence is also verbose and redundant: *"waarbij"* should be *"waarvan"*, and *"volgen"* ("follow") adds no real meaning. Moreover, the revision shifts meaning subtly but significantly—from *"het UWV betaalt"* ("the UWV pays") to *"het UWV zorgt voor de uitbetaling"* ("the UWV will handle the payment").

The revision of segment 5 is an improvement. ChatGPT clarifies that the employer should make agreements not just about the timing of the return to work, but also about the number of hours. This makes it clear that the concern is with actual working hours, not merely contractual terms. The sentence remains long, but the addition of a subheading enhances the letter's scannability.

Segment 6 contains another factual error. The sick leave benefit applies to employees who are unable to work before or after maternity leave due to pregnancy-related symptoms. The rewritten version implies the benefit is linked only to the period immediately before or after the start of the leave, which is not correct. Additionally, the paragraph ends with a sentence that does not connects properly to the preceding content.

Segment 7 opens with the somewhat awkward phrase *"de timing van de ziekte"* ("the timing of the illness"). More importantly, the revision omits central information about when the employee is entitled to sick pay and when sick days are deducted from maternity leave. Like segment 5, however, this section has been improved structurally by the addition of a heading.

Segment 8 shows improvement: the final sentence combines two original sentences into a more fluid and inviting offer to contact the organization. Finally, in segment 9, ChatGPT alters the order of the closing, placing the sender's name before the company name—reversing the convention used in the original.

	ChatGPT Simple	ChatGPT Simple (translated)
1.	Betreft: Ondersteuning bij zwangerschapsverlof en aanvragen van WAZO-uitkering	Subject: Support for maternity leave and applying for WAZO benefit
2.	Geachte heer/mevrouw,	Esteemed Sir/Madam,
3.	Gefeliciteerd met de aanstaande uitbreiding van uw team! Ik begrijp dat uw medewerker binnenkort met zwangerschapsverlof gaat. Graag informeer ik u over de stappen voor het aanvragen van de WAZO- uitkering bij het UWV, om te zorgen dat alles soepel verloopt.	Congratulations on the upcoming expansion of your team! I understand that your employee will soon be going on maternity leave. I would like to inform you about the steps to apply for the WAZO benefit through the UWV to ensure everything goes smoothly.
4.	**WAZO-uitkering aanvragen:** U kunt de zwangerschap melden en de WAZO-uitkering aanvragen via deze link. Het is belangrijk dit uiterlijk vier weken voor, maar niet later dan twee weken vóór de start van het verlof te doen. Uw medewerker	**Applying for WAZO benefit:** You can report the pregnancy and apply for the WAZO benefit through this link. It is important to do this no later than four weeks before, but no later than two weeks before, the start of the leave. Your employee

	heeft recht op minstens 16 weken verlof, waarbij minimaal 10 weken volgen na de bevalling. Tijdens het verlof verzorgt het UWV de uitbetaling van de zwangerschapsuitkering.	is entitled to at least 16 weeks of leave, whereby a minimum of 10 weeks follow after childbirth. During the leave, the UWV will handle the payment of the maternity benefit.
5.	**Plan het hervatten van werk:** Het is raadzaam om nu al te overleggen over het hervatten van het werk na het verlof, inclusief eventuele aanpassingen in de werkuren.	**Plan to resume work:** It is advisable to discuss resuming work after leave now, including any adjustments in working hours.
6.	**Aanvragen van een Ziektewetuitkering:** Mocht uw medewerker voor het ingaan van het verlof of direct erna uitvallen wegens zwangerschapsgerelateerde klachten, dan kunt u namens haar een Ziektewetuitkering aanvragen. Dit regelt u digitaal, en vermeldt u in Verzuimsignaal. Meld bij het UWV zowel de gedeeltelijke werkhervatting als volledige uitval.	**Applying for Sickness benefit:** Should your employee drop out due to pregnancy-related symptoms before the start of the leave or immediately after, you can apply for sickness benefits on her behalf. This can be done digitally, and you should report it in Verzuimsignaal. Please notify the UWV of both partial return to work and complete absence.
7.	***Let op!*** De timing van ziekte kan invloed hebben op de rechten van uw medewerker op uitkeringen. Is uw medewerker ziek binnen 6 tot 4 weken voor de vermoedelijke bevallingsdatum zonder dat het verlof is gestart, dan kan deze periode invloed hebben op de lengte van het zwangerschapsverlof. Meer details vindt u op de website van het UWV.	***Note!*** The timing of illness can affect your employee's entitlement to benefits. If your employee is sick within 6 to 4 weeks before the expected delivery date without having started her leave, this period may affect the length of her maternity leave. You can find more details on the UWV website.
8.	**Contact:** Voor vragen over de UWV-procedures kunt u contact opnemen met het UWV via 0900-9295. Voor vragen over arbodienstverlening of Verzuimsignaal sta ik u graag te woord op 035-6253147 of via s.hekker@keerpunt.nl.	**Contact:** For questions regarding UWV procedures, you can contact the UWV at 0900-9295. For questions about occupational health services or Verzuimsignaal, I am happy to assist you at 035-6253147 or via email at s.hekker@keerpunt.nl.
9.	Met vriendelijke groet, NAAM Keerpunt B.V.	With kind regards, NAME Keerpunt B.V.

Figure 5: Keerpunt letter ChatGPT Simple and translation

Compared to Ann's version, the ChatGPT Simple rewrite falls short in different areas, particularly in tone, factual accuracy, and precision. While both versions aim to improve clarity and accessibility, Ann demonstrates awareness of the professional context and the sensitivities involved in workplace communication. ChatGPT Simple, by contrast, makes several notable errors. It opens with "Congratulations on the upcoming expansion of your team," an inappropriate phrase that misrepresents the situation. It also mischaracterizes the WAZO application process by referring to "steps" when only a single action is required, and uses vague phrasing like "to ensure everything goes smoothly" without specifying what "everything" entails. More seriously, it introduces a factual error regarding the sick leave benefit, implying that it applies only immediately before or after the start of maternity leave, rather than being dependent on pregnancy-related symptoms. Key legal details—such as when sick days are deducted from maternity leave—are omitted or blurred. While both versions include subheadings and improve readability, Ann's structure is more coherent and purpose-driven. In short, ChatGPT's version lacks the contextual insight, precision, and editorial judgment that characterize Ann's professional rewrite.

When comparing the three ChatGPT prompt conditions—Simple, B1, and Process—we observe differences in both output quality and alignment with professional editing standards (see Appendix 2).

The textual comparison shows that in the B1 condition (Appendix 2), the Keerpunt letter is – in our view - optimized more than in the other ChatGPT conditions. Improvements chiefly result from simplifying language: shorter sentences, less wordiness, updated vocabulary, and more active voice. Adjustments in the B1 condition and Process condition are fairly similar. Both deliver significant improvements over the original. In the simple condition, it is noticeable that the language has hardly been simplified and in some cases has even become more complex. In addition, in the rewrite, a number of substantive things have not been included, or have been rewritten incorrectly, changing the content of the letter. Notable in this regard is that ChatGPT Simple congratulates the employer for expanding the staff. We see similar problems in the process condition (though less so). ChatGPT Process also makes errors in that rewrite, such as when it comes to who pays benefits over whom. We don't see those content problems in the B1 condition, but we do see similar problems in the Process condition.

The B1 condition outperformed the other ChatGPT prompts; this difference is explainable. Instructing the model to rewrite a text at B1 level provides a specific, well-defined goal: to simplify the language while preserving the original meaning. The B1 standard refers unambiguously to a CEFR language proficiency level, emphasizing clear and straightforward communication. In contrast, the instruction used in the *Simple* condition—to "improve the text"—is much more open-ended, leaving room for interpretation. ChatGPT may attempt to enhance style, tone, or even content, which can lead to unintended inferences and factual errors. The *Process* condition, although it includes a B1 step, involves several preceding rewriting actions during which inaccuracies may have already been introduced.

5. Results 2: quantitative comparison of human and machine rewrites

In the final stage of our investigation, we employ LiNT in a quantitative analysis to ascertain whether the revised ChatGPT texts differ from one another and from those of our editors. In total, we distinguish five groups, which we then compare on the 13 LiNT features previously mentioned. The groups or conditions are:

1. Original Texts
2. Editorial Revisions
3. ChatGPT Basic
4. ChatGPT B1
5. ChatGPT Process

To facilitate comparison, we conducted a series of Fisher's one-way analysis of variance (ANOVA) and Tukey's post hoc tests. Assumption checks indicated no meaningful deviations from normality: all varia-bles had skewness and kurtosis within acceptable ranges and Q–Q plots showed no aberrations. Homo-geneity of variances was supported by non-significant Levene's tests for all variables. The discussion is limited to variables on which we found significant effects. For convenience, we summarized all results in table 1.

Table 1. Comparison of text revisions Based on LiNT features: originals versus editor versus ChatGPT

	Originals		**Editors**		**ChatGPT Basic**		**ChatGPT B1**		**ChatGPT Process**	
Word difficulty										
Unknown words	1.85	(.25)	1.47	(.25) *	1.89	(.21) *	1.25	(.16) **	1.44	(.09)
Abstract pronouns	62.18	(12.17)	57.99	(10.57)	66.33	(8.58)	61.08	(15.34)	60.18	(14.39)
New words	88.45	(2.35)	87.04	(3.08)	93.23	(1.71) **	86.58	(5.88)	88.10	(3.68)
Sentence difficulty										
Clause length	9.88	(.84)	8.49	(.82) *	10.40	(.73) **	7.23	(.22) **	8.18	(1.25) *
Dependency length	5.45	(.68)	4.78	(.84)	5.58	(1.38)	4.90	(1.41)	5.05	(1.09)
Adjectival clauses	1.34	(.31)	.88	(.22) *	1.37	(.34) *	.57	(.17) ***	.74	(.40) **
Subordinate clauses	.31	(.07)	.37	(.11)	.40	(.12)	.28	(.06)	.25	(.08)
Enumerations	.19	(.05)	.17	(.09)	.54	(.10) ***	.64	(.32) **	.57	(.19) ***
Sentence length	12.40	(.42)	10.83	(1.09) *	13.50	(2.29) **	11.45	(2.34)	12.35	(1.82)
Text's personality										
People	130.75	(31.83)	160.50	(22.06)	114.75	(32.29) *	162.00	(23.54)	146.50	(16.74)
Personal pronouns	106.00	(18.22)	135.16	(18.15)	96.00	(32.01) *	142.25	(19.86)	130.00	(13.64)
Readers' address	62.43	(16.64)	74.70	(16.98)	52.55	(13.62)	79.30	(13.81)	73.15	(12.35)
Active verbs	90.84	(14.67)	97.80	(3.16)	95.46	(9.09)	99.17	(1.67)	99.08	(1.85)

Note. *p <.05, **p <.01, ***p <.001

5.1 Word difficulty

The initial ANOVA showed a significant effect of revision condition on the number of unknown words ($F(4, 23) = 6.81$, $p < .001$, $\eta^2 = .54$). As shown in Table 5, the ChatGPT B1 group had the lowest mean ($M = 1.25$, $SD = 0.16$), while ChatGPT Simple and the original texts had the highest ($M = 1.89$, $SD = 0.21$; $M = 1.85$, $SD = 0.25$). Tukey's HSD tests confirmed that ChatGPT B1 had significantly fewer unknown words than both ChatGPT Simple ($p < .01$) and the original texts ($p < .05$). Editor texts also had significantly fewer unknown words than ChatGPT Simple and the original ($p < .05$). Additionally, ChatGPT Process differed significantly from Simple ($p = .008$). These results suggest that both ChatGPT B1 and editor revisions effectively simplified the text, performing comparably. No significant group differences were found for abstract pronouns ($p = .81$) or new words ($p = .056$).

5.2 Sentence difficulty

An ANOVA showed a significant effect of condition on clause length ($F(4, 23) = 9.87$, $p < .001$, $\eta^2 = .63$). ChatGPT B1 produced the shortest clauses ($M = 7.23$, $SD = 0.22$), while ChatGPT Simple and the original texts had the longest ($M = 10.40$, $SD = 0.74$; $M = 9.88$, $SD = 0.84$). Tukey post hoc tests found that clause length in ChatGPT B1 and ChatGPT Process was significantly shorter than in ChatGPT Simple ($p < .001$ and $.008$). ChatGPT B1 also differed significantly from the original texts ($p < .01$), though ChatGPT Process did not. Editor texts had shorter clauses than ChatGPT Simple ($p < .05$). Overall, editors and ChatGPT B1 were most effective in simplifying sentence length.

A significant group difference also emerged in adjectival clause use ($F(4, 23) = 7.13$, $p < .001$, $\eta^2 = .55$). ChatGPT B1 had the lowest incidence ($M = 0.57$, $SD = 0.17$), while ChatGPT Simple and the original texts had the highest ($M = 1.37$, $SD = 0.34$; $M = 1.34$, $SD = 0.31$). Tukey tests confirmed ChatGPT B1 used significantly fewer adjectival clauses than both ChatGPT Simple and the original texts. Editor texts also used fewer adjectival clauses than ChatGPT Simple ($p < .05$). ChatGPT Process showed similar reductions compared to both ChatGPT Simple ($p = .026$) and the original ($p = .016$).

Clause enumeration also varied significantly by condition ($F(4, 23) = 12.10$, $p < .001$, $\eta^2 = .68$). ChatGPT B1 had the highest average ($M = 0.64$, $SD = 0.32$), followed by ChatGPT Process ($M = 0.57$, $SD = 0.19$) and ChatGPT Simple ($M = 0.54$, $SD = 0.10$). Original texts and editor versions had lower means ($M = 0.19$, $SD = 0.05$; $M = 0.17$, $SD = 0.09$). All ChatGPT variants showed significantly more enumerations than both the original and editor texts ($p < .001$). ChatGPT Process also showed a significant improvement over the original ($p = .016$), though differences among the ChatGPT outputs were not significant.

5.3 The text's personality

An ANOVA revealed a significant effect of revision condition on references to people ($F(4, 23) = 3.40$, $p = .025$, $\eta^2 = .37$). ChatGPT B1 ($M = 162.00$, $SD = 23.54$) and editor texts ($M = 160.50$, SD = 22.06) had higher scores than ChatGPT Simple (M = 114.75, SD = 32.29) and the original texts (M = 130.75, SD = 31.83). Tukey's tests confirmed that both ChatGPT B1 and editor texts included significantly more references to people than ChatGPT Simple ($p < .05$). ChatGPT Process did not differ significantly from any other condition.

An ANOVA also showed significant differences in personal pronoun usage ($F(4, 23) = 4.47$, $p = .008$, $\eta^2 = .44$). ChatGPT B1 (M = 142.25, SD = 19.86) and editor texts (M = 135.17, SD = 18.15) used more personal pronouns than ChatGPT Simple (M = 96.00, SD = 32.01) and the original texts (M = 106.00, SD = 18.22). Post hoc analysis showed ChatGPT B1 differed significantly from ChatGPT Simple ($p < .01$), and editor texts also differed from ChatGPT Simple ($p < .05$). No significant differences involved ChatGPT Process. Finally, the variables "new words" and "sentence length" showed marginal effects: new words ($F(4, 23) = 2.69$, $p = .056$) and sentence length ($F(4, 23) = 2.66$, $p = .059$). Other variables did not reach statistical significance.

5.4 Conclusion on rewriting modalities

Based on the LiNT analysis, *ChatGPT B1* appears to be the most effective rewriting mode. It produces the lowest scores for word and sentence difficulty, uses enumerations appropriately, and maintains a personal tone through references to people and personal pronouns. As such, B1 is well-suited for rewriting texts for broad audiences, improving readability and accessibility while preserving the original meaning. Rewrites by professional editors show similar performance, particularly in sentence simplification and vocabulary clarity. Both approaches—*ChatGPT B1* and human editing—clearly outperform both the original texts and the *ChatGPT Simple* condition. While the *ChatGPT Process* mode uses more enumerations, which can support comprehension, it also introduces more content errors and inconsistencies. A striking observation is the underperformance of *ChatGPT Simple*. Despite its goal of simplification, this mode results in longer sentences, more adjectival clauses, and a higher occurrence of unfamiliar words. It also includes fewer personal pronouns and references to people, reducing engagement and relatability.

In short, *ChatGPT Simple* struggles to balance simplicity and clarity. In contrast, both the *B1* mode and human editors deliver more readable, consistent, and personalized rewrites. Notably, only these two also produced texts free of factual errors. While *ChatGPT Process* shows some strengths in structure, it falls short on several key dimensions. Overall, *ChatGPT B1* proves to be an efficient and reliable rewriting method, delivering results comparable to those of professional editors.

6. Limitations and conclusions

Like any study, this research has limitations. First, it focused on a small set of text types, which limits the generalizability of the findings. Rewriting requirements can vary widely across genres—such as scientific articles, news releases, instructions, or marketing materials—and the effectiveness of revision strategies likely depends on text type. Future research should explore a broader range of genres.

Second, while the study examined 13 LiNT features related to readability and accessibility, these do not capture all dimensions of text quality. Stylistic nuance, tone consistency, and alignment with communicative goals were addressed only partially in the qualitative analysis. As such, the LiNT measures, though useful for targeted quantitative analysis, are insufficient for evaluating the overall quality of more complex documents, such as reports or policy papers.

The study also relied on a small group of editors from a single agency, which may have limited stylistic diversity. Their shared background could influence the uniformity of their rewrites, reducing the representativeness of professional variation. Additional research involving editors with diverse training and experience would help address this. Moreover, the editors worked under time constraints and without direct input from a client—conditions that differ from real-world practice, where access to context and goals typically leads to higher-quality output.

On the AI side, ChatGPT's responses are generated probabilistically, introducing variation between outputs. In this study, responses were generated in a single attempt ("one-shots"), without revision. Different runs may have yielded different results, which affects the reproducibility of findings.

Finally, the study did not examine collaborative workflows between editors and ChatGPT. Human and AI rewrites were produced in isolation, yet in practice, ChatGPT could support editors by simplifying sentences, generating structure, or reducing lexical difficulty. Human refinement of AI drafts would likely lead to better outcomes than either working alone. Similarly, editor performance in this study may have been improved with access to client input, which is standard in professional contexts.

Our findings demonstrate that ChatGPT, as a generative AI tool, can be a valuable asset in professional rewriting processes, especially when used under the guidance of experienced writers or editors. A comparative analysis of texts rewritten by professional editors and ChatGPT shows that both approaches can lead to improved readability and comprehensibility. However, important differences remain between the qualitative, context-sensitive approach of human editors and ChatGPT's more algorithmic, prompt-based method.

Quantitative analyses indicate that human editors achieved significant improvements in readability. Their revisions featured simpler vocabulary, shorter sentences, more active constructions, and increased use of personal pronouns. These features made the texts better suited to their intended audience by reducing complexity and enhancing accessibility. ChatGPT also demonstrated the ability to improve readability, particularly when prompted with clear and targeted instructions. However, the quality of its output was inconsistent, and inaccuracies or misleading formulations were not uncommon.

Qualitative analysis of the rewriting processes showed that human editors paid greater attention to tone, style, and context. They not only ensured technical correctness but also added structure through headings and reading aids, eliminated vague phrasing, and occasionally enriched the text with additional service elements, such as extra contact information or clarifications. Human editors also excelled at handling context-specific challenges, such as personalizing tone or omitting irrelevant content. In contrast, ChatGPT sometimes introduced errors or misinterpretations, particularly when dealing with content that required contextual sensitivity.

The analysis further showed that human editors displayed greater flexibility in adapting their strategies to the specific text and target audience. ChatGPT, by contrast, was limited by the instructions it received and lacked the ability to adapt beyond the scope of the prompt. These findings underline the importance of human control and editorial oversight when integrating generative AI into professional writing workflows.

6.1 What is next for pedagogy and practice?

The rise of AI—particularly generative models like ChatGPT—raises important questions about the future of writing skills, pedagogy and practice. Traditionally, writing has been seen as a distinctly human activity, requiring creativity, critical thinking, and language proficiency. But as technology evolves, so too does our understanding of what writing entails. This is not a new debate: the introduction of tools like the word processor and the Internet has long shaped how we define writing competence and what aspects can be supported—or replaced—by technology.

We now face a more radical shift. Whereas tools like navigation systems made map-reading obsolete for drivers, chatbots like ChatGPT can produce entire texts independently. This development presents challenges for domains such as education, where student performance is often assessed through writing. While we do not pursue that discussion here, it underscores a broader concern: What should humans still be able to do, and what can be delegated to AI?

At its core, writing is the ability to express ideas, information, and arguments clearly and effectively through language. Before AI, this process—from planning to final editing—was carried out entirely by humans. Tools like ChatGPT now automate parts of that process, leading to a shift in what it means to be a skilled writer. Where writing was once defined by the ability to create a coherent text from a blank page, emerging skillsets include the ability to prompt, guide, and critically assess AI-generated content. Writing, in this view, becomes more akin to "prompt engineering"—directing the AI to achieve a desired outcome. This raises fundamental questions about whether writing proficiency is still defined by independent creation, or by effective use of technological tools to achieve communicative goals.

Our findings reinforce a central point for writing *pedagogy*: using generative AI productively and responsibly hinges on the very same knowledge and skills that traditional writers need. Effective prompting is not a technical trick; it is applied rhetorical practice that requires clarity about purpose, audience, genre, constraints, and text quality criteria. Likewise, evaluating and revising AI output demands the same expert reading and judgment writers bring to their own drafts. In short, generative AI amplifies existing writerly knowledge rather than replacing it (Flower & Hayes, 1981; Hayes, 2012).

Prompt design maps closely onto the rhetorical situation: writers must specify communicative goals, audience characteristics, genre conventions, and constraints on style and register. A growing body of work in technical and professional communication conceptualizes prompt engineering as a rhetorical activity that centers audience and purpose (Ranade, Saravia, & Johri, 2025). Classic document-design research reminds us that quality also depends on reader-focused choices (Schriver, 1997), while language-proficiency frameworks such as the CEFR help instructors make audience/readability targets explicit in prompts and evaluations (Council of Europe, 2020).

Despite AI's ability to generate fluent text, its outputs still contain errors, inconsistencies, and a lack of contextual depth or critical reasoning. This underscores the continued importance of human skills—particularly critical thinking, contextual judgment, and editorial oversight. Writing may evolve into a hybrid competence, in which the human role shifts from sole author to evaluator, adaptor, and corrector of AI-generated drafts.

Assessing AI-generated prose engages the same competencies used to evaluate human writing: rhetorical fit, genre alignment, coherence and cohesion, evidence quality, and correctness. Decades of writing research show that high-quality revision is guided by purpose- and audience-sensitive diagnosis, not surface editing (Sommers, 1980; Faigley & Witte, 1981). Contemporary outcomes frameworks (e.g., the WPA Outcomes Statement) similarly emphasize rhetorical knowledge, critical reading, and process-based revision—competencies that transfer directly to supervising AI (Council of Writing Program Administrators, 2014). Framed this way, "AI literacy" for writers is best understood as an extension of existing writing literacies (Long & Magerko, 2020).

Human–AI collaboration research indicates that expertise strongly conditions outcomes. When tasks fall within AI's strengths and users exercise judgment, quality and productivity can rise substantially; when they fall outside that "jagged technological frontier," performance can degrade, especially without expert oversight (Dell'Acqua et al., 2023). Novices are also more susceptible to automation bias and over-reliance on incorrect algorithmic suggestions, whereas domain expertise helps users detect and correct such errors (Dratsch et al., 2023; Romeo & Conti, 2025; Gaube et al., 2021). These patterns support a pedagogical stance that the best users of AI are skilled writers and critical readers—and that using AI without those skills does not reliably build them. Writing courses should – in our view - therefore treat AI as a studio tool that extends—never replaces—core writing instruction. The best users of AI are skilled writers and critical readers. GenAI can accelerate drafting and broaden options, but it does not eliminate the need for writerly knowledge. Pedagogy should center that knowledge and make AI supervision an explicit, assessable part of how we teach composing and revision.

The AI evolution also reinforces a broader shift in how writing is practiced: from an individual to a collaborative activity. While collaboration in writing is not new, AI introduces a new type of partner—the machine. With models like ChatGPT capable of generating content, structures, and revisions, the human writer increasingly takes on the role of curator or editor rather than original author. Writers become managers of an interactive process: generating drafts through AI and refining them to meet communicative goals.

This dynamic is comparable to writing from sources, where writers gather and repurpose existing material (albeit now with much greater ease and scale). For professional editors, this introduces a more complex form of collaboration—not just with clients, but also with generative AI (see figure . Clients themselves may use AI, and may or may not involve external editors. The division of labor among client, editor, and AI remains unclear (cf. Schriver 2012), and how this relationship evolves depends largely on the capabilities and limitations of AI systems.

As this study shows, relying on GenAI alone—especially by non-professionals—entails risks. Not every prompt yields appropriate or accurate output. Human oversight remains essential in ensuring quality, reliability, and contextual appropriateness in AI-assisted writing.

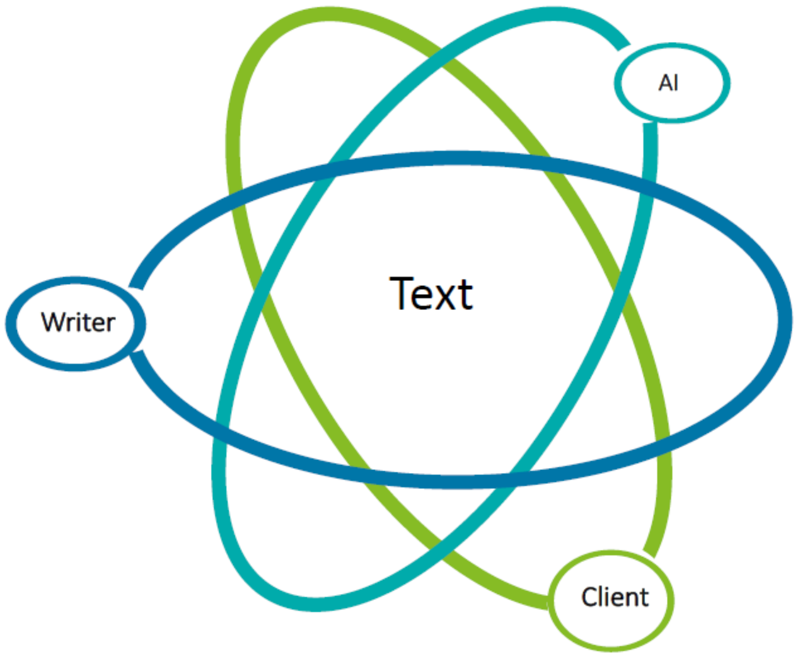

Figure 6: Collaboration in GenAI fueled professional writing

Our outlook is more optimistic than that of the translators and writers referenced in the introduction. In organizational contexts, texts are often so context-specific that AI is unlikely to produce them autonomously without professional oversight. Research also shows that human–AI collaboration is most effective when it builds on human strengths. Vaccaro et al. (2024) found that combining human and AI input leads to performance gains when humans outperform AI—but not the other way around. This suggests that professional writers can benefit from AI support, while lay users may see fewer advantages.

AI also has the potential to streamline writing processes (Noy & Zhang, 2023), allowing professionals to focus more on content and strategy. Editors in our study spent an average of 13 minutes per task; with an AI-generated draft, this could likely be reduced. We are currently exploring this in follow-up studies.

At the same time, increased reliance on AI poses risks. It may reduce the motivation to develop critical writing skills or cultivate a personal voice, especially if a prompt can generate a usable draft. In this sense, AI may lead to both efficiency gains and a loss of valuable abilities—particularly in writing education.

Generative AI presents both opportunities and challenges. It will likely play a central role in future writing, but human input—creativity, critical thinking, and judgment—will remain essential. Like past innovations such as the typewriter and word processor, generative AI is already reshaping how we write.

References

Alkaissi, H., & McFarlane, S. I. (2023). Artificial hallucinations in ChatGPT: implications in scientific writing. Cureus, 15(2). 0.7759/cureus.35179

Anthropic. (2023, July 11). Claude 2. https://www.anthropic.com/news/claude-2 Anthropic

Anzidei, Christopher M. (2002) The Revision Process in Legal Writing: Seeing Better to Write Better. Legal Writing 23, 8 (2002), 23-57.

Arabas, C. (1990). Technical Writing in a Corporate Culture; A Study of the Nature of Information. New Jersey: Sablex.

Benites, F., Delorme Benites, A. & Anson, C. M. (2023). Automatic Text Generation and Summarization for Academic Writing. Kruse, O. et al. (eds), Digital Writing Technologies in Higher Education, Springer. https://doi.org/10.1007/978-3-031-36033-6.

Bisaillon, J. (2007). Professional Editing Strategies Used by Six Editors. Written Communication, 24(4), 295-322. https://doi.org/10.1177/0741088307305977

Brynjolfsson, E., Li, D., & Raymond, L. R. (2023). Generative AI at work (NBER Working Paper No. 31161). National Bureau of Economic Research. https://doi.org/10.3386/w31161

Buçinca, Z., Malaya, M. B., & Gajos, K. Z. (2021). To trust or to think: Cognitive forcing functions can reduce overreliance on AI in AI-assisted decision-making. Proceedings of the ACM on Human-Computer Interaction, 5(CSCW1), Article 188. https://doi.org/10.1145/3449287

Burrough-Boenisch, J. (2006). Negotiable Acceptability: Reflections on the Interactions between Language Professionals in Europe and NNS 1 Scientists Wishing to Publish in English. Current Issues in Language Planning, 7(1), 31–43. https://doi.org/10.2167/cilp086.0

Calderhead, J. (1981. Stimulated Recall: a Method for Research on Teaching. British Journal of Educational Psychology 51(2):211-217. https://doi.org/10.1111/j.2044-8279.1981.tb02474.x

Council of Europe. (2020). Common European Framework of Reference for Languages: Learning, teaching, assessment—Companion volume. Council of Europe Publishing. https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4

Council of Writing Program Administrators. (2014). WPA Outcomes Statement for First-Year Composition (Version 3.0). https://wpacouncil.org/aws/CWPA/pt/sp/statements

Cross, G. A. (1990). A Bakhtinian exploration of factors affecting the collaborative writing of an executive letter of an annual report. Research in the Teaching of English, 24(2), 173-203.

Dell'Acqua, F., et al. (2023). Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality. Harvard Business School Technology & Operations Management Working Paper (24-013). https://doi.org/10.2139/ssrn.4573321

Dratsch, T., Chen, X., Rezazade Mehrizi, M. H., Mähringer-Kunz, A., Püsken, M., Baeßler, B., Sauer, S., Maintz, D., Pinto dos Santos, D., & Kloeckner, R. (2023). Automation bias in mammography: The impact of AI BI-RADS suggestions on reader performance. Radiology, 307(4), e222176. https://doi.org/10.1148/radiol.222176

Duin, A. H. (1991). Computer-Supported Collaborative Writing: The Workplace and the Writing Classroom. Journal of Business and Technical Communication, 5(2), 123-150. https://doi.org/10.1177/1050651991005002001

Ede, L and Lunsford, A. (1990). Singular Texts /Plural Authors: Perspectives on Collaborative Writing. Carbondale. Southern Illinois Press.

Faigley, L., & Witte, S. (1981). Analyzing revision. College Composition and Communication, 32(4), 400–414. https://doi.org/10.2307/357848

Flower, L., & Hayes, J. R. (1981). A Cognitive Process Theory of Writing. College Composition and Communication, 32(4), 365–387. https://doi.org/10.2307/356600

Gaube, S., Suresh, H., Raue, M., Merritt, A., Berkowitz, S. J., Lermer, E., Coughlin, J. F., Guttag, J. V., Colak, E., & Ghassemi, M. (2021). Do as AI say: Susceptibility in deployment of clinical decision-aids. NPJ Digital Medicine, 4, 31. https://doi.org/10.1038/s41746-021-00385-9

Google. (2023, August 29). Write emails, documents and more with ease and speed using Duet AI in Google Docs and Gmail. Google Workspace Updates. Workspace Updates Blog

Google. (2023, July 13). Bard's latest update: More features, languages and countries. https://blog.google/products/gemini/google-bard-new-features-update-july-2023/

Graham, Steve. (2018). A Revised Writer(s)-Within-Community Model of Writing. Educational Psychologist. 53. 1-22. 10.1080/00461520.2018.1481406.

Hayes, J. R. (2012). Modeling and Remodeling Writing. Written Communication, 29(3), 369-388. https://doi.org/10.1177/0741088312451260

Hayes, J. R., Flower, L., Schriver, K. A., Stratman, J. F., & Carey, L. (1987). Cognitive processes in revision. In S. Rosenberg (Ed.), Advances in applied psycholinguistics, Vol. 2. Reading, writing, and language learning (pp. 176–240). Cambridge NJ: Cambridge University Press.

Heilmann, Till. A. ((2023). The Beginning of Word Processing; A Historical Account. Kruse, O. et al. (eds), Digital Writing Technologies in Higher Education, Springer. https://doi.org/10.1007/978-3-031-36033-6.

Herbold, S., Hautli-Janisz, A., Heuer, U., Kikteva, Z., & Trautsch, A. (2023). A large-scale comparison of human-written versus ChatGPT-generated essays. Scientific Reports, 13 (18617). https://doi.org/10.1038/s41598-023-45644-9

Horning, A. (2006). Professional writers and revision. Revision: History, theory, and practice, 117-141.

Janssen, D. M. L., & Schilperoord, J. (1992). Policy writing research and functional text quality. In H. Pander Maat, & M. Steehouder (Eds.), Studies in functional text quality. Utrecht Studies in Language and Communication. Amsterdam: Rodophi.

Janssen, D. M. L., & van der Mast, N. P. (2001). Collaborative writing for the government. In D. M. I. Janssen, & R. Neutelings (Eds.), Reading and writing public documents (pp. 169-208). Amsterdam: John Benjamins.

Janssen, D.M.L. (1991). Schrijven aan beleidsnota's (Writing policy papers). PhD thesis. Groningen: Wolters Noordhoff.

Janssen, D.M.L., & Neutelings, R. (2001). Reading and writing public documents. Amsterdam: John Benjamins.

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

Kleijn S. (2018). Clozing in on readability: How linguistic features affect and predict text comprehension and on-line processing. PhD-disseration. Utrecht: LOT.

Lay, M. and Karin, W.M. (1991). Collaborative Writing in Industry: Investigations in Theory and Practice. Amityville NJ: Baywood.

Lee, M., Gero, K. I., Chung, J. J. Y., et al. (2024). A design space for intelligent and interactive writing assistants. CHI '24. https://dl.acm.org/doi/10.1145/3613904.3642697

Lee, M., Liang, P., & Yang, Q. (2022). CoAuthor: Designing a human-AI collaborative writing dataset for exploring language model capabilities. CHI '22. https://arxiv.org/abs/2201.06796

Leijten, M. (2007). Writing and Speech Recognition: Observing Error Correction Strategies of Profes-sional Writers. Utrecht: LOT.

Leijten, M., & Van Waes, L. (2013). Keystroke logging in writing research: Using Inputlog to analyze and visualize writing processes. Written Communication, 30(3), 358-392. https://doi.org/10.1177/0741088313491692.

Leijten, M., Waes, L. V., Schriver, K., & Hayes, J. R. (2014). Writing in the workplace: Constructing documents using multiple digital sources. Journal of Writing Research, 5(3), 285–337. DOI: 10.17239/jowr-2014.05.03.3

Lewis, P., Perez, E., Piktus, A., et al. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. NeurIPS. https://arxiv.org/abs/2005.11401

Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 1–16. https://doi.org/10.1145/3313831.3376727

Longpre, S., Hou, L., Vu, T., et al. (2023). The Flan collection: Designing data and methods for effective instruction tuning. ICML. https://proceedings.mlr.press/v202/longpre23a.html

Lundstrom, K. & Baker, W. (2009). To give is better than to receive: The benefits of peer review to the reviewer's own writing, Journal of Second Language Writing 18 (1), 30-43. https://doi.org/10.1016/j.jslw.2008.06.002.

Mistral AI. (2023a, September 27). Announcing Mistral 7B. https://mistral.ai/news/announcing-mistral-7b

Mistral AI. (2023b, December 11). Mixtral of experts (Mixtral 8×7B). https://mistral.ai/news/mixtral-of-experts

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. Science, 381, 187 - 192. https://doi.org/10.1126/science.adh2586.

OpenAI. (2023a, March 1). Introducing ChatGPT and Whisper APIs (forum announcement). https://community.openai.com/t/introducing-chatgpt-and-whisper-apis/80485

OpenAI. (2023b, May 18). Introducing the ChatGPT app for iOS. https://openai.com/index/introducing-the-chatgpt-app-for-ios/

Ouyang, L., Wu, J., Jiang, X., et al. (2022). Training language models to follow instructions with human feedback. NeurIPS. https://arxiv.org/abs/2203.02155

Pander Maat, H. and M. Steehouder (Eds.), Studies in functional text quality. Utrecht Studies in Language and Communication. Amsterdam: Rodophi.

Pander Maat, H., & Dekker, N. (2016). Tekstgenres analyseren op lexicale complexiteit met T-Scan. Tijdschrift voor Taalbeheersing, 38(3), 263-304. https://doi.org/10.5117/TVT2016.3.PAND

Ranade, N., Saravia, M., & Johri, A. (2025). Using rhetorical strategies to design prompts: A human-in-the-loop approach to make AI useful. AI & Society, 40, 711–732. https://doi.org/10.1007/s00146-024-01905-3

Reuters. (2023, February 1). ChatGPT sets record for fastest-growing user base. https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/

Reynolds, F.J. et al. (1995). Professional writing in context : lessons from teaching and consulting in worlds of work. Hillsdale NJ: Erlbaum.

Romeo, G., & Conti, D. (2025). Exploring automation bias in human–AI collaboration: A review and implications for explainable AI. AI & Society https://doi.org/10.1007/s00146-025-02422-7

Schriver, K. (2012). What do technical communicators need to know about information design? J. Johnson-Eilola & S. Selber, Eds. Solving problems in technical communication, pp. 386–427. University of Chicago Press.

Schriver, K. A. (1997). Dynamics in document design: Creating text for readers. Wiley.

Selzer, J. (1983). The Composing Processes of an Engineer. College Composition and Communication, 34(2), 178–187. https://doi.org/10.2307/357405

Selzer, J. (1984). Exploring Options in Composing. College Composition and Communication, 35(3), 276–284. https://doi.org/10.2307/357456

Shibani, Antonette (2023), Analytic Techniques for Automated Analysis of Writing. Kruse, O. et al. (eds), Digital Writing Technologies in Higher Education, Springer.

Sommers, N. (1980). Revision strategies of student writers and experienced adult writers. College Composition and Communication, 31(4), 378–388. https://doi.org/10.2307/356588.

Vaccaro, M., Almaatouq, A. & Malone, T. (2024) When combinations of humans and AI are useful: A systematic review and meta-analysis. Nature Human. Published online.

Van der Mast, N.P. and Janssen, D.M.L. (2001). Strategic revision of public documents: how policy writers revise documents in order to reach consensus. In D.M.L. Janssen, & R. Neutelings (Eds.), Reading and writing public documents (pp. 209-230). Amsterdam: John Benjamins.

Van Riel, C. B. and Fombrun, C. J. (2007). Essentials of Corporate Communication: Implementing Practices for Effective Reputation Management. (n.p.): Taylor & Francis.

Van Steendam, E. (2016). Editorial: Forms of collaboration in writing. Journal of Writing Research, 8(2), 183–204. https://doi.org/10.17239/jowr-2016.08.02.01

Van Waes, L. (1991). Rapporteren met computer en pen & papier: Verschillen in het pauze- en revisiegedrag bij ervaren schrijvers. Toegepaste taalwetenschap in artikelen, 40(1), 176-190. https://doi.org/10.1075/ttwia.40.17wae

Wiggers, K. (2023, July 11). Anthropic releases Claude 2, its second-gen AI chatbot. TechCrunch. https://techcrunch.com/2023/07/11/anthropic-releases-claude-2-the-second-generation-of-its-ai-chatbot/ Alkaissi, H., & McFarlane, S. I. (2023). Artificial hallucinations in ChatGPT: implications in scientific writing. Cureus, 15(2). 0.7759/cureus.35179

Willey, I., & Tanimoto, K. (2012). "Convenience Editing" in action: Comparing English teachers' and medical professionals' revisions of a medical abstract. English for Specific Purposes, 31(4), 249-260. https://doi.org/10.1016/j.esp.2012.04.001

Yao, S., Zhao, J., Yu, D., et al. (2022). ReAct: Synergizing reasoning and acting in language models. ICLR. https://arxiv.org/abs/2210.03629

Yuan, A., Coenen, A., Reif, E., Ippolito, D., & Davis, L. (2022). Wordcraft: Story writing with large language models. IUI '22. https://dl.acm.org/doi/10.1145/3490099.3511105

Zhai, X., Chu, X., Wang, M., Tsai, C. C., Liang, J. C., & Spector, J. M. (2024). A systematic review of Stimulated Recall (SR) in educational research from 2012 to 2022. Humanities and Social Sciences Communications, 11(1), 1-14. https://doi.org/10.1057/s41599-024-02987-6

Zhang, Z., Gao, J., Dhaliwal, R. S., & Li, T. J.-J. (2023). VISAR: A human-AI argumentative writing assistant with visual programming and rapid draft prototyping. UIST '23. https://toby.li/files/uist23-zhang-visar.pdf