Korean Journal of English Language and Linguistics, Vol 24, December 2024, pp. 1415-1435
DOI: 10.15738/kjell.24..202412.1415

KOREAN JOURNAL OF ENGLISH LANGUAGE AND LINGUISTICS

ISSN: 1598-1398 / e-ISSN 2586-7474

http://journal.kasell.or.kr

South Korean STEM Graduate Students' Use of ChatGPT in Self-Initiated L2 Writing: A Process-Tracing Study*

Siwon Sung · In Chull Jang (Seoul National University)

Received: November 11, 2024
Revised: December 22, 2024
Accepted: December 26, 2024

Sung, Siwon (First author)
Master student, Department of English Language Education,
Seoul National University
Email: siwonsung98@snu.ac.kr

Jang, In Chull (Corresponding author)
Professor, Department of English Language Education,
Seoul National University
Email: icjang@snu.ac.kr

*The article is based on the first author's MA thesis.

ABSTRACT

Sung, Siwon and In Chull Jang. 2024. South Korean STEM graduate students' use of ChatGPT in self-initiated L2 writing: A process-tracing study. *Korean Journal of English Language and Linguistics* 24, 1415-1435.

Using process-tracing technologies such as webcam-based eye-tracking, screen recording, and stimulated recall, this study explored how South Korean STEM graduate students employed ChatGPT in their self-initiated and naturalistic English writing settings. The analysis of three 30-minute writing sessions and stimulated recalls, as well as semi-structured interviews from four participants, revealed the complex and dynamic patterns of tool use and behaviors in ChatGPT-mediated writing. At the tool level, participants employed a range of digital tools alongside ChatGPT, guided by personal preferences and writing community norms. At the behavioral level, they demonstrated that using ChatGPT in L2 writing involved not only diverse gaze and non-gaze behaviors but also individualized sequences of these behaviors to meet specific writing goals. The closer analysis of the two participants' cases showed idiosyncratic patterns in tool usage and behavior. One participant used ChatGPT in a brief and targeted way, focusing only on specific vocabulary and expression searches. The other adapted ChatGPT-related behaviors dynamically based on the purpose and stage of writing. By unobtrusively tracing L2 writers' real-time evolutions during the writing process, this study suggests that ChatGPT has urged us to rethink conventional writing stages and behaviors and to develop new research methods that can capture such dynamic processes of AI-mediated L2 writing.

KEYWORDS

academic writing, CALL, ChatGPT, eye-tracking, writing process

1. Introduction

The development of Artificial Intelligence (AI) has made profound impacts on perceptions, motivations, and methodologies of second language education. The public release of ChatGPT has led to a range of discussions regarding the objectives of language education, the roles of teachers, and the ethics of plagiarism, particularly since it generates texts in response to a prompt, unlike other new technologies such as machine translators. ChatGPT's ability to generate text has challenged traditional writing practices, which have long been viewed as complex textual and discursive processes that require intensive cognitive involvement and conventions of a discourse community. As such, Warschauer et al. (2023) point out that new writing practices enabled by AI tools will be "as transformational as was the original transition to digital writing" (p. 1). Similarly, Godwin-Jones (2022) argues that AI tools will not only support writers throughout the conventional stages of the writing process but "reshape the process of authoring and editing" (p. 5). While the expansion of ChatGPT is expected to significantly transform conventional writing literacy (Godwin-Jones 2022, Warschauer et al. 2023), it remains unclear how this new AI tool might affect a writer's actions in actual writing processes. To address this gap, this study explores micro-processes of tool use and behavioral patterns in ChatGPT-mediated L2 writing, focusing on South Korean graduate students' self-initiated academic writing tasks.

Although ChatGPT has gained significant attention from L2 researchers for its innovative performance and potential as a writing reference tool, the majority of empirical studies have focused on investigating language learners' perceptions of using ChatGPT in completing their writing tasks or proposing methods of pedagogical interventions and examining their effects on language learning (Yan 2023, Zou and Huang 2023). What such studies have not fully addressed is that most digital tools can be used not only in classrooms where teachers regulate students' use for pedagogical purposes but also in out-of-class settings, or what Sauro and Zourou (2019) call "the digital wilds." Similarly, ChatGPT can be used in everyday writing environments where an individual student's agency, regulations, and goals play important roles in successful writing. Thus, the context of this study is not within the classroom but in a naturalistic setting where participants have the flexibility to use various tools and behave in ways they usually do.

To investigate writing processes in naturalistic settings, this study draws attention to data collection methods for tracing writing processes. As Hellmich and Vinall (2023) note, process-tracing methodologies "allow researchers to empirically observe how participants interact in technologically-mediated spaces and with technological tools" (p. 3). For example, screen recording technologies have enabled researchers to capture writing processes without their intrusion and increase ecological validity (Séror and Gentil 2023). Along with screen recording, this study employed an innovative technology, webcam-based eye-tracking, to capture a finer level of micro-level writing processes through gaze movement and fixation.

The remainder of this paper consists of the following sections. First, it reviews the literature on the use of ChatGPT and process-tracing studies in L2 writing. Then, the research contexts, participants, data collection procedures, and analytic methods are provided. Specifically, the method section introduces the coding scheme of tools and behaviors that were discovered and categorized from the process-tracing data. The results section presents an overview of the general tool use and behaviors of four participants in ChatGPT-mediated writing, followed by a detailed analysis of two participants' cases. Finally, this paper discusses the characteristics of ChatGPT-mediated writing in a naturalistic setting as well as the methodological and pedagogical implications for future research.

2. Literature Review

2.1 ChatGPT and L2 Writing

ChatGPT, developed by OpenAI, is an artificial intelligence-generated content (AIGC) model capable of understanding natural language and generating contextually appropriate responses to user prompts in an interactive manner. Unlike traditional Computer-Assisted Language Learning (CALL) tools, which entail fixed functions and predefined educational goals, ChatGPT is a general-purpose tool offering a wide range of possibilities for L2 writing support. Since its capabilities are not solely determined by its technological affordance but are shaped through its interactions with the user, ChatGPT can produce highly adaptable and responsive results to users' specific input. While such inherent flexibility of ChatGPT in use and outcomes may offer personalized support to L2 writers' immediate needs, it also introduces unpredictable and individualized usage patterns.

Although research on ChatGPT's role in L2 writing remains in its early stages, the majority of existing studies focus on exploratory discussions of ChatGPT's affordances as a writing assistant and writers' perceptions regarding its use in L2 writing. Since ChatGPT's public release and recognition of its innovative functionalities, scholars in L2 writing and CALL have begun discussing its potential benefits and challenges in L2 writing contexts (Barrot 2023). For instance, Su et al. (2023) demonstrated its ability to provide scaffolded instructions and feedback to assist students' argumentative writing. However, Warschauer et al. (2023) point out that ChatGPT might trigger contradictions that could ultimately have adverse effects on EFL writers.

Several intervention studies have explored how EFL writers perceive ChatGPT during writing tasks. Zou and Huang (2023) examined ChatGPT's perceived roles in L2 writing by gathering reflective journals and conducting focus group interviews with 219 doctoral candidates at a public university in China. They found that most students viewed ChatGPT as a "personal tutor" or a "learning partner." However, participants also expressed concerns about potential drawbacks, such as learning loss, dilution of authorial voice, the generation of unintelligent text, and risks to academic integrity. Using survey data from 339 university students, Chan and Hu (2023) examined students' perceptions of ChatGPT's influence on their writing. They found that participants generally held positive views and expressed familiarity with its effects on their writing processes. Similarly, Liu and Ma (2024) surveyed 405 EFL learners in China to examine their perceptions of ChatGPT outside the classroom. They found that although perceived ease of use did not affect learners' attitudes toward ChatGPT, positive attitudes toward the tool significantly increased learners' intention to use it in informal digital learning contexts.

While existing studies provide significant insights into language learners' perceptions of this new tool for L2 writing, they often overlook the actual behaviors and interactions that are essential for understanding its impact on the L2 writing process. Rather than examining ChatGPT use in educator- or researcher-structured writing contexts where students must follow predefined workflows, it is important to uncover how L2 writers utilize ChatGPT in their everyday writing tasks within authentic real-world contexts. These writing environments can reveal the personalized and spontaneous ways in which learners integrate ChatGPT into their authentic writing practices. Thus, this study aims to explore the naturalistic moment-by-moment aspects of L2 writers' ChatGPT use and discern the writing behaviors that emerge in ChatGPT-mediated L2 writing processes.

2.2 Process-Tracing Studies in L2 Writing

There has been a notable increase in the use of innovative methodologies within L2 writing process research (Manchón and Roca De Larios 2023, Kessler 2024). Process-tracing technologies such as screen capture

technology, keystroke logging, and eye-tracking are being actively employed for methodological triangulation, allowing for detailed observation and analysis of L2 writers' digital composing processes (Latif 2019, Leijten and Van Waes 2013, Manchón and Roca De Larios 2023). In traditional writing process studies, methods such as interviews, think-aloud protocols, and stimulated recalls were commonly utilized. While these methods provide valuable insights into writers' perceptions and intentions during the writing process, their heavy reliance on participants' self-reports restricted researchers from directly investigating behaviors and strategies that occurred during writing (Michel et al. 2020). To address these methodological limitations, direct observation methods such as keystroke logging and screen recording technology emerged as alternatives (Hamel et al. 2015, Révész et al. 2019). While used in conjunction with verbal reports, these methods have proven to be minimally invasive yet highly effective, allowing for real-time capture of actions on writers' digital screens during composition. However, one limitation in L2 writing process research is that these methods do not sufficiently capture the locations and flows of writers' visual attention. As Leijten et al. (2014) poignantly point out, gaze behaviors matter in the writing process as reading is a prerequisite for reviewing and planning.

Recent scholars are triangulating screen capture technology and keystroke logging with eye-tracking data to address these issues. Gánem-Gutiérrez and Gilmore (2018) investigated the temporal aspects of writing using screen recording, eye-tracking, and stimulated recall with 22 Japanese EFL writers. They collected gaze plot recordings during a 35-minute essay-writing task. Their results indicated that participants dedicated the most time to text construction and revision, frequently involving re-reading. Similarly, Liu and Yu (2022) used eye-tracking, stimulated recall, and reflective journals to explore how 24 Chinese EFL writers engaged with different feedback types (direct and indirect) in an automated writing evaluation (AWE) system. They collected 30 minutes of eye gaze plot recordings while participants revised their essays using the AWE system. Based on 25 online behaviors identified from gaze plot recordings, their study revealed that participants invested more time and cognitive resources in processing indirect feedback. Collectively, these studies illustrate how "eye-tracking methodology is diversifying and breaking boundaries, [with] its innovation fueled by different disciplinary traditions" (Godfroid et al. 2020, p. 245). By capturing writers' temporal and dynamic interactions with multiple resources in digital writing contexts, this methodology can provide new insights into the complexities of the writing process.

Although eye-tracking technology offers significant methodological advantages, it has not been widely used in L2 writing process research due to its high cost and limited accessibility outside laboratory settings (Stickler and Shi 2017). In response to these logistical constraints, webcam-based eye-tracking is emerging as a practical alternative (Yang and Krajbich 2021). This technology enables observations in naturalistic environments by leveraging webcams commonly integrated into laptops. Although webcam-based eye-tracking does not achieve the same level of accuracy and precision as traditional eye-tracking technology, it can still provide valuable insights into L2 writing processes. When triangulated with screen recording and stimulated recall data, webcam-based eye-tracking has the potential to enhance the ecological validity of research, offering a practical solution for studying authentic writing behaviors in real-world contexts.

3. Method

3.1 Participants and Writing Tasks

This study selected STEM graduate students as participants because English writing plays an important role in the academic communities of science (Kim 2020). Researchers and graduate students in the STEM areas are

usually required to write their research outputs in English, which motivates them to engage with L2 writing in authentic settings. Of the seven potential participants recruited four South Korean STEM graduate students from three universities in Seoul, South Korea, were selected through purposeful sampling. The selection criteria targeted individuals who regularly used ChatGPT in their writing and were actively working on academic writing tasks as part of their ongoing coursework or research projects. Table 1 presents the profiles of the four participants.

Table 1. Participant Profiles

Name	Degree, Field of Study	Gender, Age	Self-Initiated Writing Task	English Proficiency	Number of Publications	ChatGPT Experience
Sam	M.S., Smart Vehicle Engineering	Male, 27	Journal Article	Upper-Intermediate (B2) TOEIC 850	K: 2 E: 1	6 months
Jake	Ph.D., AI	Male, 27	Journal Article	Advanced (C1) NEW TEPS 493	K: 0 E: 9	2 years (Since GPT 2.0)
Collin	M.S., AI	Male, 26	Thesis	Advanced (C1) TOEFL 110	K: 0 E: 4	8 months
Thomas	Ph.D., AI	Male, 29	Research Proposal	Upper-Intermediate (B2) NEW TEPS 330 TOEIC 790	K: 0 E: 2	3 months

Note. K = the number of publications in Korean, E = the number of publications in English. All names are pseudonyms. Participants' CEFR levels were determined by aligning their official certification test scores with the CEFR.

The researcher did not assign specific writing tasks to observe participants' ChatGPT use during their authentic writing process. Instead, participants were asked to engage in self-initiated writing tasks already in progress or planned. The participants also chose the places and times for writing, facilitated by the mobility of webcam-based eye-tracking. This research context aimed to examine participants' writing as "a phenomenon situated in context" (Lafford 2009, p. 674) and ensure ecological validity by observing their naturalistic writing behaviors during real-world tasks.

3.2 Data Collection

To understand participants' use of ChatGPT in a naturalistic setting, this study collected multiple types of data, including semi-structured interviews (2 hours per participant), gaze plot recordings (90 minutes per participant), and stimulated recall interviews (approximately 3 hours per participant). Two semi-structured interviews were conducted at the beginning and end of the research. In the first interview, participants were asked about their language learning backgrounds, experiences, and perceptions of L2 writing. The second interview was conducted to obtain information on participants' experiences with and evaluations of ChatGPT during writing for this project.

The gaze plot recordings and subsequent stimulated recalls served as the primary data source of this study. The gaze plot recordings were collected through a webcam-based eye-tracking technology, RealEye (https://www.realeye.io/). This technology enabled the authors to capture participants' writing processes in "natural, non-controlled, non-laboratory conditions" (Séror and Gentil 2023, p. 142). The data was in video format, where screen recordings were overlaid with gaze data visualized through nodes and lines, capturing participants' real-time on-screen activities and "spatial distribution and duration of eye fixation" (Godfroid 2019, p. 237). Figure 1 presents a screenshot of a gaze plot recording from one of the participants. The gaze data, represented by yellow nodes and connecting lines, illustrates the movement of the participant's gaze pattern across the screen, where two software applications—Dynalist on the left and ChatGPT on the right—are displayed in a split-screen

configuration (See Appendix A for details about each software application). In writing process research, gaze plot recordings offer enhanced insights compared to standard screen recordings, particularly by capturing behaviors during periods of viewing that lack corresponding mouse movements or typing actions (Gánem-Gutiérrez and Gilmore 2018, Michel et al. 2020). Gaze data provides detailed information about participants' reading behaviors during writing, their navigation across diverse semiotic and digital resources, and the duration and frequency of their attention to specific textual elements. This methodology enables researchers to conduct detailed analysis of micro-processes of writing, such as reading, scanning, spell-checking, and consulting various external resources (Bailey and Withers 2018, Takayoshi 2016). In this study, participants completed three 30-minute eye-tracking sessions while engaging in self-initiated writing tasks. Before these sessions, participants underwent a tutorial to familiarize themselves with the webcam-based eye-tracking tool and completed a trial run.

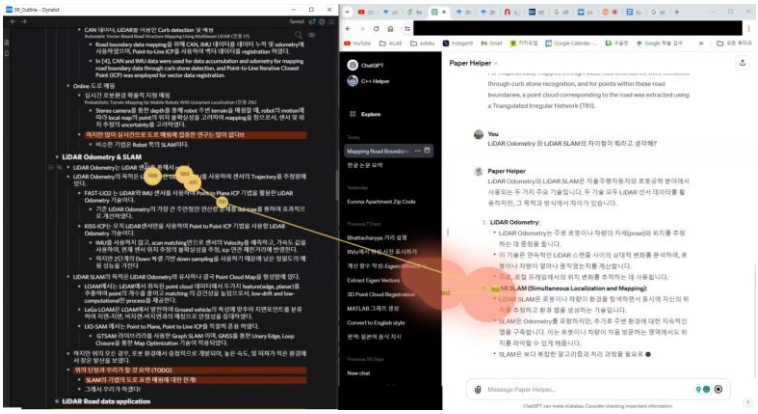


Figure 1. Screenshot of a Gaze Plot Recording

To triangulate the gaze plot recordings, stimulated recalls were employed to gather participants' verbal accounts of their use of and intentions with ChatGPT in the writing processes. A stimulated recall was conducted and recorded via Zoom within three days of each session. While watching the entire 30-minute gaze plot recording obtained through webcam-based eye-tracking, participants self-reported their intentions and cognitive processes related to their online behaviors and interactions with ChatGPT. Figure 2 illustrates the data collection process.

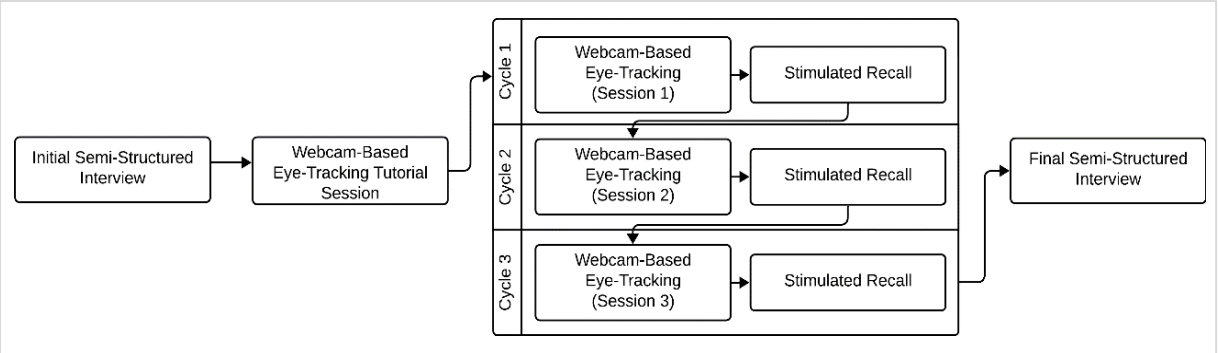


Figure 2. Data Collection Process

3.3 Analysis

The data analysis was conducted in three phases to explore participants' writing practices with ChatGPT. Figure 3 presents a flow chart of the data analysis process.

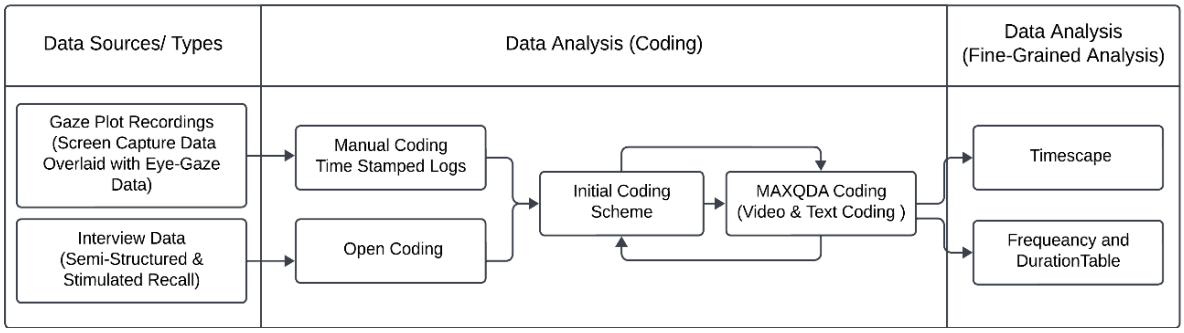


Figure 3. Data Analysis Process

3.3.1 Phase 1: Time-stamped logs and open coding

Before each stimulated recall session, participants' gaze plot recordings were analyzed and converted into time-stamped logs to provide an overall understanding of the writing process throughout each 30-minute eye-tracking session. Moment-to-moment events were documented in a time-stamped log, including language-related episodes, ChatGPT consultations, eye-gaze movements, software arrangements, and relevant contextual details. Additionally, the areas were marked where the eye-tracking data appeared ambiguous or the intentions behind certain behaviors were unclear. During the stimulated recalls, participants were asked to clarify these parts and review the behaviors identified in the time-stamped logs. Subsequently, the stimulated recall data and semi-structured interviews were transcribed and open-coded (Strauss and Corbin 1998). An initial coding scheme was established through an iterative process of comparing data from the gaze plot recordings, time-stamped logs, stimulated recalls, and semi-structured interviews.

3.3.2 Phase 2: Video coding

Following a preliminary analysis of the participants' writing processes in the first phase, a fine-grained analysis of gaze plot recordings was conducted using the video coding function in MAXQDA 24, as shown in Figure 4. The coding was performed on two levels: tool and behavior. The tool level refers to the word processing or reference software that participants used during writing, such as ChatGPT, Overleaf, and Google. On the other hand, the behavior level represents participants' micro-processes during composition, including text construction, ChatGPT prompting, reading their own text, arranging tools, zig-zagging gaze movements, and copying and pasting. After video coding, participants' responses from the stimulated recalls were tagged with each corresponding tool or behavior code using MAXQDA's memo function to corroborate the labeled coding themes. Following the initial coding of all gaze plot recordings, two additional cycles of video coding were conducted to ensure consistency. The final coding schemes are presented in Appendices A and B.

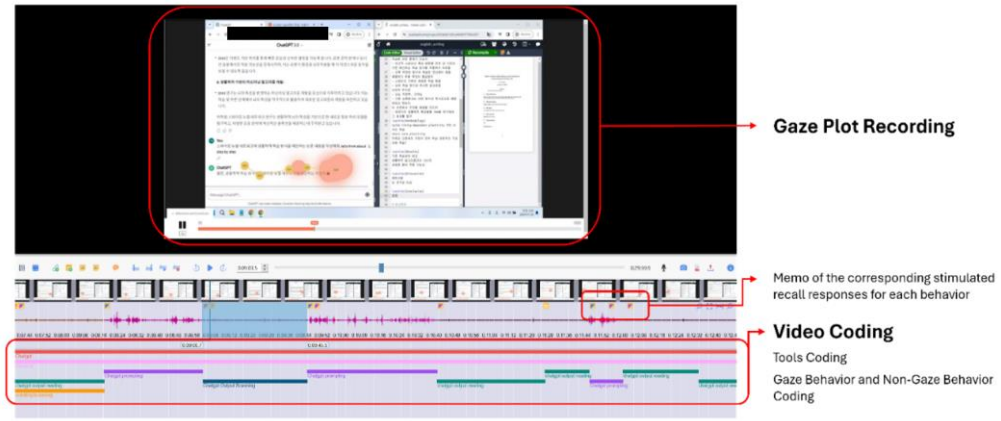


Figure 4. Video Coding by MAXQDA

3.3.3 Phase 3: Timescape and frequency/duration table

In this phase, the coded segment data from MAXQDA were exported into a spreadsheet to create timescapes and frequency/duration tables for comparative analysis between participants. Timescape, initially proposed by Smith (2017) to explore L2 writers' modal movements during multimodal composition tasks, is a method for visualizing the real-time composing process. This study employed timescapes to identify general patterns and individual differences in behavioral evolution during writing. Using the code coverage data from coded segment data, the sequence and proportion of time participants allocated to each tool and behavior were represented in two separate bars, with each code assigned a different color (see Figures 5 and 6 for examples). Additionally, tables were created to indicate the temporal distribution of each code at both the tool and behavior levels. Referring to Gánem-Gutiérrez and Gilmore (2018) and Roca de Larios et al. (2008), the duration and frequency of each code were summed to illustrate how much time participants dedicated to each tool or behavior (see Tables 4 and 5 for examples).

4. Findings

To illustrate the emergent micro-processes of participants' engagement with ChatGPT during writing, this section first presents the general patterns of tool use and behavior identified through qualitative coding. It reveals that ChatGPT-mediated academic writing in a naturalistic setting is entangled with various tools and moment-by-moment relevant behaviors. Then, the real-time writing processes of two participants, Collin and Thomas, are analyzed using a set of timescape and frequency/duration tables based on the tool and behavior coding schemes. These cases demonstrate how the interplay of tools and behaviors evolves throughout the writing sessions in greater detail.

4.1 Tools and Behaviors in ChatGPT-Mediated Writing

In a naturalistic setting, ChatGPT-mediated writing was characterized by the use of various software tools and the entanglement of multiple behaviors involving these tools. While completing self-initiated writing tasks,

participants used not only word processing software but also reference tools such as ChatGPT, search engines, and automated writing evaluation tools they found helpful. Table 2 presents the tools used by the four participants across their three writing sessions. Although the frequency and duration of tool use varied according to each participant's writing goals and context, the types and specifications of tools reflect both their writing communities and their strategies for managing complex academic writing processes in digital spaces. For instance, all four participants used Overleaf for writing because of its powerful LaTeX editing and collaboration features. They stated that most natural sciences and engineering researchers use this software due to journal requirements for LaTeX formatting and its ability to integrate references, codes, and figures directly into the manuscript. This integration also streamlines collaborative writing, which is common in the STEM fields.

Table 2. Overview of Participants' Tool Use

Participants	Sam			Jake			Collin			Thomas		
Session / Tool Code	S1	S2	S3	S1	S2	S3	S1	S2	S3	S1	S2	S3
ChatGPT	*	*	*	*	*	*	*	*	*	*	*	*
Overleaf	*	*	*			*	*	*	*	*	*	*
Google			*		*	*	*	*	*	*		
Dynalist	*	*	*									
Text Document Comparison Site											*	*
ArXivGPT						*						
Notion				*	*	*						
Others: Naver Dictionary, PDF		*			*	*	*	*	*			
Grammarly	*	*					*	*	*			

Table 3 outlines participants' behaviors while using ChatGPT in their writing tasks. These behaviors demonstrated that digital writing comprises complex micro-processes involving not only text production but also visual processing, such as (re)reading one's own or source texts (Hayes 1996) and navigating external online resources (Leijten et al. 2014). Based on the nature of digital writing (Gánem-Gutiérrez and Gilmore 2018, Leijten et al. 2014, van Weijen et al. 2009) and the affordances of gaze-plot recording data, the participants' behaviors in ChatGPT-mediated writing were categorized into *gaze behaviors* and *non-gaze behaviors*.

This study identified six types of gaze behaviors: *ChatGPT output scanning*, *ChatGPT output reading*, *re-reading, zig-zagging, Overleaf compile and reading or scanning*, and *software scanning* (See Appendix B for details). Before consulting ChatGPT, participants often began by re-reading or scanning their texts written in their L1 or L2. After prompting ChatGPT, they employed various strategies to process its response. Participants sometimes thoroughly read the output, fixating on all or most words (*ChatGPT output reading*). In contrast, in others, they quickly scanned the output, moving their gaze horizontally or focusing only on targeted text segments (*ChatGPT output scanning*). A *zig-zagging* gaze pattern was common as participants shifted their gaze between ChatGPT and other software arranged side-by-side on a split screen. This behavior allowed them to compare, cross-reference, or integrate elements from their original text with ChatGPT's responses.

For non-gaze behaviors, 11 types were identified, as shown in Table 3 (See Appendix B for details). These primarily involved typing or mouse movements. For example, *ChatGPT prompting* referred to inputting text into ChatGPT, either by typing or pasting text composed beforehand. While some participants invested time crafting precise prompts, others simply copied and pasted text without detailed prompt engineering. *Incorporating*

ChatGPT's output into writing involved transferring full or selected responses into another software, either by copying and pasting or typing manually. This process was often followed by *revising ChatGPT output* to align with the writer's intent or to incorporate field-specific terminology. As participants used various tools throughout the writing process, their methods of transferring text across software—such as *drag and drop* or manual typing—also influenced their ChatGPT use. Most participants preferred dragging and dropping text from their writing software into ChatGPT. After generating an output, they incorporated it into their writing either by dragging and dropping or typing manually, reflecting individual preferences shaped by their attitudes toward ChatGPT.

These tools and behaviors highlight the idiosyncratic yet contextually shaped dynamics of ChatGPT-mediated writing. The following section will focus on two participants, Collin and Thomas, to explore their tool use and behaviors in greater depth, illustrating when, where, and how they used ChatGPT throughout their writing process.

Table 3. Overview of Participants' Gaze and Non-Gaze Behaviors

	Participants	Sam			Jake			Collin			Thomas		
Behavior Category	Session / Behavior Code	S1	S2	S3	S1	S2	S3	S1	S2	S3	S1	S2	S3
Gaze Behavior (GB)	ChatGPT output scanning	*	*	*	*	*	*	*	*	*	*	*	*
	ChatGPT output reading	*	*	*	*	*	*	*	*		*	*	*
	Zig-zagging	*	*	*							*	*	*
	Re-reading	*	*	*		*	*	*	*	*		*	*
	Overleaf compile and reading or scanning	*	*	*			*	*	*	*	*		*
	Software scanning	*	*	*	*	*	*	*	*	*	*	*	*
Non-gaze Behavior (NGB)	ChatGPT prompting	*	*	*	*	*	*	*	*	*	*	*	*
	Text construction (K)	*	*	*							*	*	
	Text construction (E)		*		*		*				*		*
	Revising writing	*	*	*		*	*	*	*	*			*
	Drag and drop	*	*	*		*	*	*	*	*	*	*	*
	Incorporating ChatGPT output into writing	*	*	*	*	*	*	*	*	*	*	*	*
	Revising ChatGPT output	*	*	*			*						*
	Search		*	*		*	*	*	*	*	*	*	
	Locating	*	*	*	*	*		*		*	*	*	*
	Grammarly suggestion apply	*	*					*	*	*			
	Other	*	*	*	*		*				*	*	*

4.2 Case 1: Collin

4.2.1 Session Overview

At the time of data collection, Collin was a master's student specializing in AI whose primary motivation for English writing stemmed from his thesis work. His English proficiency exceeded most engineering students at his

institution, evidenced by his track record of four published English-language articles. Having used ChatGPT since May 2023, Collin has accumulated approximately eight months of experience with the tool.

During three eye-tracking sessions, Collin worked on a self-initiated task: adapting his published journal article into his thesis. He employed three distinct strategies during this transformation. First, he conducted comprehensive paraphrasing at the word, sentence, and paragraph levels. Second, he reintegrated detailed contents omitted to meet journal publication length constraints. Third, he incorporated new arguments and insights derived from conference presentation feedback. Throughout these sessions, Collin utilized ChatGPT as what he termed an "expression dictionary" in one of his stimulated recalls. He employed it as a linguistic resource for identifying academically appropriate phrases and expressions, enabling him to refine his writing while maintaining a scholarly register.

4.2.2 Tools

Figure 5 presents a timescape illustrating the pattern of tool use across Collin's three eye-tracking sessions. Overleaf served as his primary word-processing platform for revision and review. The data revealed intermittent and brief engagements with both ChatGPT and Google. As the single line in Figure 5 indicates, Collin typically utilized one tool at a time, switching between tabs when consulting different resources. Rather than arranging ChatGPT and Overleaf in parallel view, he accessed ChatGPT only when needed, returning to the Overleaf tab after each consultation.

Figure 5. Collin' Tool Timescape

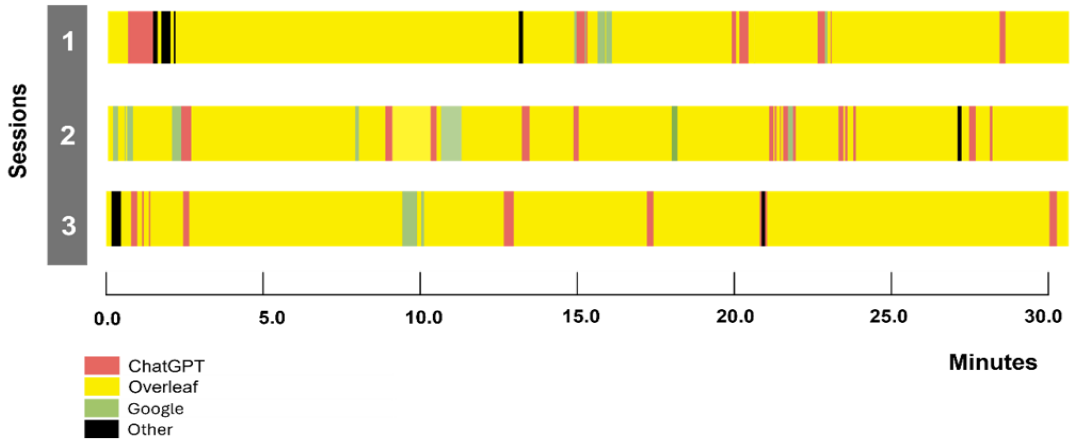


Table 4 summarizes the frequency and duration of digital tool use across all three sessions. The data aligns with the patterns visualized in the timescape. Overleaf dominated tool usage, averaging approximately 26 minutes and 50 seconds per 30-minute session. In contrast, ChatGPT averaged 1 minute and 48 seconds per session, while Google averaged 41 seconds. However, despite their brief total durations, both ChatGPT and Google were accessed frequently. Notably, ChatGPT was accessed 9 times in session 1 (totaling 1 minute and 54 seconds) and 15 times in session 2 (totaling 2 minutes and 11 seconds). This pattern of frequent, brief interactions suggests that Collin's ChatGPT use was highly targeted and purpose-specific—a phenomenon examined in detail in the following section.

Table 4. Collin's Tool Code Frequency and Duration

Tool Codes	Session 1		Session 2		Session 3	
	f	Duration	*f*	Duration	*f*	Duration
ChatGPT	**9**	**01:54.2**	**15**	**02:11.0**	9	01:19.8
Google	5	00:36.8	6	00:55.5	2	00:32.0
Overleaf	14	26:51.6	22	25:57.7	12	27:44.3
Others	4	00:15.9	3	00:55.9	2	00:23.9

Note. *f* = Frequency.

4.2.3 Behaviors

Figure 6 visualizes the distinct behavioral patterns that emerged across the three ChatGPT-mediated writing sessions. The timescape revealed two primary patterns: 1) dominance of *re-reading* and *revising writing* and 2) a sequence of behaviors centered on *ChatGPT prompting*. As represented by green and light-yellow straps on the timescape, Collin dedicated significant time to *re-reading* and *revising writing* in Overleaf. Between these dominant activities is a series of thin straps representing multiple behaviors occurring within a short timeframe. These behaviors typically followed a sequence: *revising writing*, *drag and drop*, *ChatGPT prompting*, *ChatGPT output scanning*, *incorporating ChatGPT output into writing*, *revising writing*, and *Overleaf compile and reading*. Notably, the *ChatGPT prompting* strap often appears imperceptible due to its brevity; Collin typically copied and pasted revised text into ChatGPT without explicit instructions, quickly scanned the output, returned to Overleaf, manually typed selected suggestions, adjusted surrounding text for coherence, and reviewed the compiled document.

The timescape also revealed white-colored straps (indicating Google searches) adjacent to ChatGPT-related behaviors. Stimulated recall data indicates that Collin employed these searches either to explore alternatives before using ChatGPT or to verify the appropriateness of ChatGPT's suggestions prior to incorporation.

Figure 6. Collin' Behavior Timescape

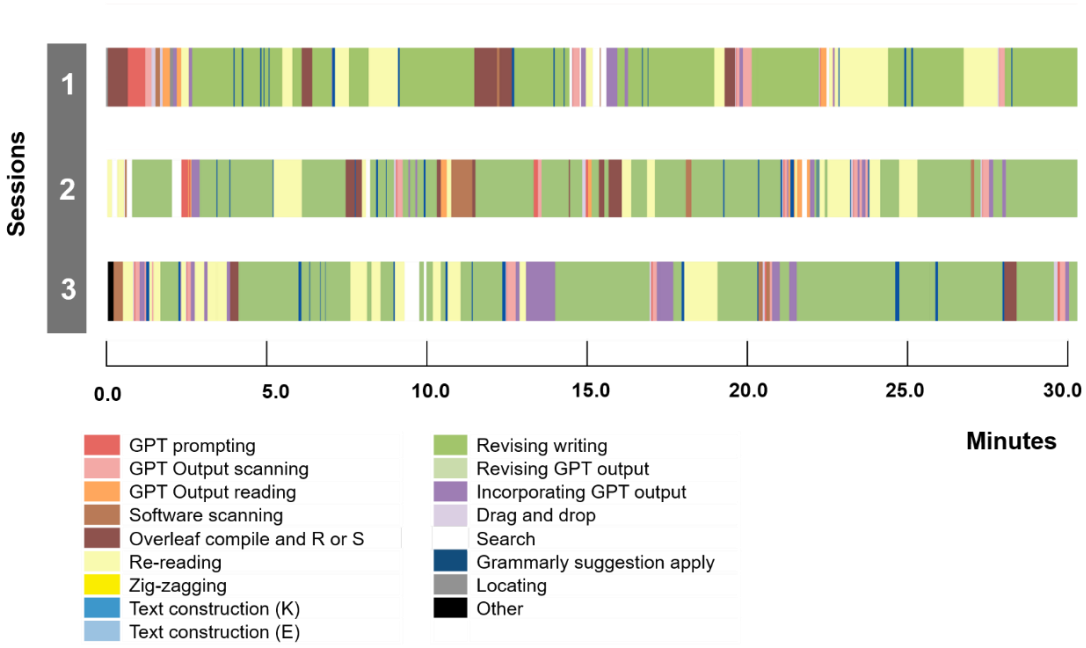


Table 5 shows the duration and frequency of these behaviors across the writing sessions. *Revising writing* dominated time allocation (averaging 17 minutes per session), followed by *re-reading* (averaging 4 minutes per session). In contrast, ChatGPT-related behaviors—*ChatGPT output scanning* and *ChatGPT prompting*—each consumed less than one minute per session. Session 2 exemplifies the rapidity of these interactions: *ChatGPT prompting* occurred eight times within 33 seconds (averaging 6.6 seconds per prompt), while *ChatGPT output scanning* comprised seven instances over 50 seconds (averaging 7.1 seconds per scan). This pattern of brief, focused interactions suggests that Collin employed it in a restrained manner rather than utilizing ChatGPT's diverse affordances. In a simulated recall, he stated that he tried to use ChatGPT specifically for limited purposes of searching academic phrases.

Table 5. Collin's Behavior Code Frequency and Duration

Behavior Codes		Session 1		Session 2		Session 3	
		f	Duration	*f*	Duration	*f*	Duration
GB	ChatGPT output scanning	6	00:57.8	**7**	**00:50.4**	8	00:59.3
	ChatGPT output reading	3	00:33.0	8	0:00:42		
	Re-reading	10	05:12.2	12	03:51.6	14	04:36.7
	Overleaf compile and reading or scanning	5	02:21.0	7	01:18.8	2	00:38.1
	Software scanning	3	00:14.1	4	0:01:00	3	00:32.0
NGB	ChatGPT prompting	5	00:36.3	**8**	**00:33.6**	6	00:13.6
	Revising writing	24	17:26.3	30	18:57.4	23	18:27.7
	Drag and drop	6	00:18.6	5	00:19.1	6	00:18.6
	Incorporating ChatGPT output into writing	8	00:57.9	9	00:56.7	12	02:37.0
	Search	5	00:36.2	6	00:56.0	2	00:32.3
	Locating	2	00:09.2			1	00:10.6
	Grammarly suggestion apply	16	00:37.5	15	00:35.2	15	01:09.5

4.3 Case 2: Thomas

4.3.1 Session Overview

Thomas was a doctoral student majoring in AI, with an upper-intermediate level of English proficiency and experience publishing two research articles in English. At the time of data collection, he had been using ChatGPT for writing assistance for approximately three months. He participated in this study with the self-initiated goal of drafting a research proposal; he aimed to develop his research ideas into a three-to-four-page document to discuss with his supervisor. Unlike Collin, who maintained a consistent writing purpose across sessions, Thomas's goals varied in each session, leading him to engage with ChatGPT in diverse ways. In the first session, he used ChatGPT to brainstorm research ideas and create an outline in Korean. In the following session, he elaborated on this outline, adding specific details to each subsection he had established in session one. During the final session, he translated and expanded his Korean outline into an English manuscript, line by line, beginning with the introduction. Throughout this process, he used ChatGPT to find suitable vocabulary and expressions and to revise the draft at the end.

4.3.2 Tools

Figure 7 visually represents Thomas's use of various software tools across the three writing sessions. Notably, each session includes two (and sometimes three) layers of lines, indicating that Thomas often used multiple software tools simultaneously. His routine involved setting up a split view with ChatGPT and Overleaf. He frequently adjusted this configuration to include additional reference tools, such as a Text Document Comparison Site, particularly in the second and third sessions. He used this comparison tool to highlight differences between his original writing and the output refined by ChatGPT, as it visually marked changes with color.

Figure 7. Thomas' Tool Timescape

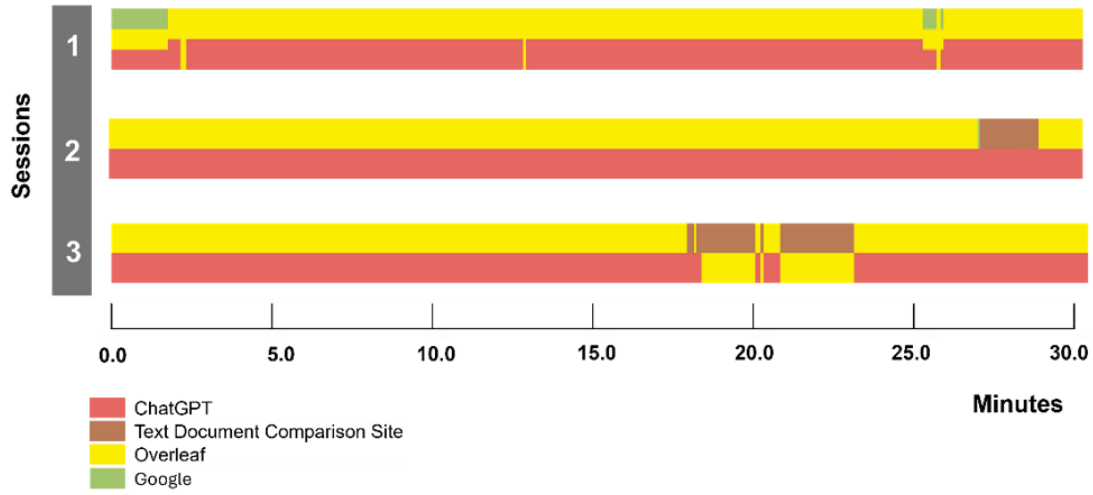


Table 6 further illustrates that Thomas had multiple software tools open while writing, with the concurrent use of Overleaf and ChatGPT dominating his screen time. On average, he spent 26 minutes out of each 30-minute session using this ChatGPT-Overleaf setup. Additionally, his reliance on the Text Document Comparison Site increased as his writing focus shifted from brainstorming and planning to drafting, reflecting how his tool usage adapted to his evolving writing goals across the sessions.

Table 6. Thomas' Tool Code Frequency and Duration

Tool Codes	Session 1		Session 2		Session 3	
	f	Duration	*f*	Duration	*f*	Duration
ChatGPT + Overleaf	4	27:21.9	2	28:07.2	5	25:36.1
ChatGPT + Text Document Comparison Site			**1**	**01:49.0**	**2**	**00:23.3**
Overleaf + Text Document Comparison Site					**3**	**04:00.6**
ChatGPT + Google + Overleaf	3	02:15.6				
Overleaf	3	00:22.5				
ChatGPT + Google			1	00:03.8		

4.3.3 Behaviors

Figure 8 illustrates how different behaviors unfold across three sessions of ChatGPT-mediated writing. The most noticeable feature of Thomas's behavior timescape is the variation in visual patterns across sessions, compared to Collin's more consistent pattern (see Figure 6). This variation reflects Thomas's distinct purposes in each session: brainstorming and outlining in Korean, elaborating on the outline in Korean, and drafting the outline in English. These evolving purposes significantly influenced how he employed ChatGPT throughout the writing process.

In the first session, which focused on brainstorming and outlining his research proposal, *ChatGPT prompting* and *ChatGPT output scanning* were the primary activities that recurred frequently. To facilitate brainstorming, Thomas asked ChatGPT questions related to his field, quickly scanned the responses, and prompted new questions based on the ideas generated. In the second session, Thomas mainly engaged in *ChatGPT prompting* and *ChatGPT output reading*, often followed by an iterative sequence of *zig-zagging* and *incorporating ChatGPT output into writing*. Unlike the first session, where *ChatGPT output scanning* was dominant, he was willing to thoroughly read ChatGPT's responses to expand his outline while *zig-zagging* between tools to incorporate relevant content into Overleaf.

In the third session, Thomas's ChatGPT use became more intermittent and focused, following a pattern of *ChatGPT prompting*, *ChatGPT output reading* or *scanning*, and selectively *incorporating ChatGPT output into writing*. This shift aligns with the session's primary goal—to draft in English. Specifically, as he translated his Korean outline into English, Thomas consulted ChatGPT to find appropriate expressions and terminology. He then prompted ChatGPT to revise his English version, carefully comparing the output with his original text. Finally, he incorporated selected parts into Overleaf by manually typing or copying and pasting.

Figure 8. Thomas's Behavior Timescape

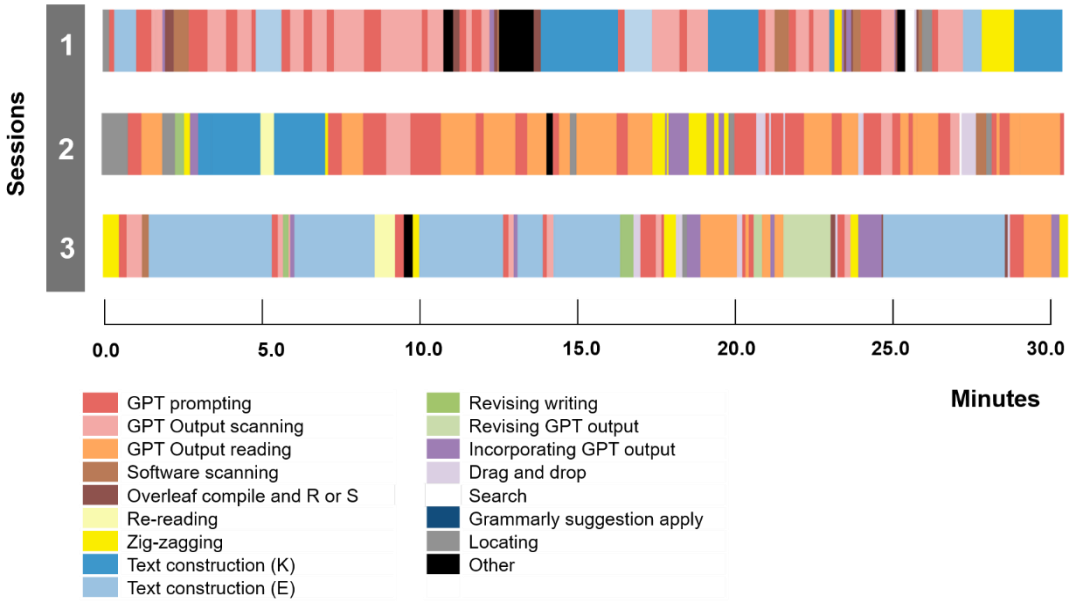


Table 7 shows that the frequency and duration of behaviors generally align with the ChatGPT usage patterns observed in the timescape. *ChatGPT prompting* was more frequent and sustained during the brainstorming and

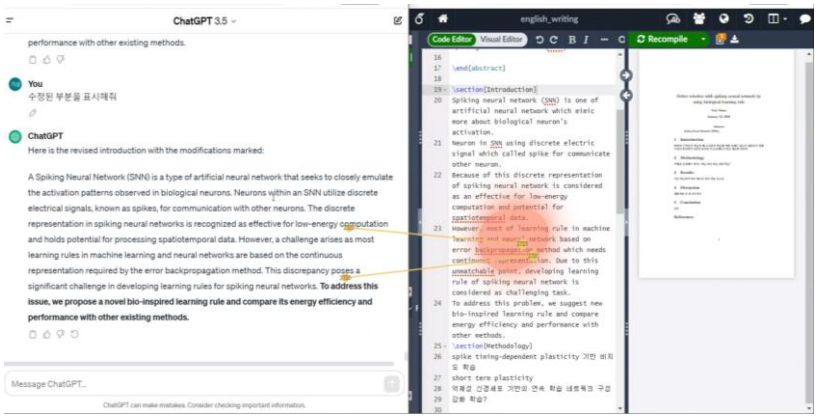
outlining stages of sessions 1 and 2. In session 1, *ChatGPT prompting* totaled 5 minutes and 30 seconds (19 instances); in session 2, it increased to 7 minutes and 37 seconds (20 instances); but in session 3, it dropped significantly to 2 minutes and 25 seconds (11 instances). This variation suggests that Thomas relied more heavily on ChatGPT during the brainstorming phase and adopted a more independent approach during drafting.

Another notable behavior shown in Table 6 is *zig-zagging*. This gaze pattern, consistent with Thomas's habit of displaying multiple tools on one screen, appeared across all three sessions, averaging 1 minute and 20 seconds. This duration closely aligns with the time spent *incorporating ChatGPT output into writing* in sessions 2 and 3, highlighting their co-occurrence during the outlining and drafting phases, as illustrated in the timescape. For instance, in session 2, *zig-zagging* lasted 1 minute and 32 seconds, while *incorporating ChatGPT output into writing* took 1 minute and 39 seconds. This suggests that Thomas accepted and integrated ChatGPT's suggestions into his writing after continuous verification and careful consideration.

Table 7. Thomas's Behavior Code Frequency and Duration

	Behavior Codes	Session 1		Session 2		Session 3	
		f	Duration	*f*	Duration	*f*	Duration
GB	ChatGPT output scanning	18	09:16.9	3	01:23.8	7	01:24.2
	ChatGPT output reading	2	01:40.1	16	10:28.7	5	02:59.3
	Zig-zagging	2	01:14.3	**7**	**01:32.7**	5	01:34.3
	Re-reading			1	00:25.4	1	00:37.8
	Overleaf compile and reading or scanning	9	01:02.6			3	00:16.6
	Software scanning	5	01:17.3	1	00:18.9	1	00:12.6
NGB	ChatGPT prompting	**19**	**05:30.6**	**20**	**07:37.5**	**11**	**02:25.3**
	Text construction (K)	4	05:40.7	3	03:31.9		
	Text construction (E)	2	01:16.8			6	15:20.2
	Revising writing					2	00:35.7
	Drag and drop	1	00:04.1	5	01:00.1	7	00:43.7
	Incorporating ChatGPT output into writing	5	00:29.2	**6**	**01:39.0**	7	01:57.6
	Revising ChatGPT output					1	01:27.8
	Search	1	00:17.3	1	00:03.9		
	Locating	2	00:31.8	5	01:46.0	1	00:08.6
	Grammarly suggestion apply						
	Others	3	01:38.4	1	00:12.1	1	00:16.3

Figure 9. Thomas's Zig-Zagging Behavior



5. Discussion

This qualitative study has examined how South Korean STEM graduate students employed ChatGPT in their self-initiated writing processes. Unlike existing empirical studies, which typically focus on students' perceptions and teacher interventions, this project analyzed L2 writers' moment-to-moment interactions with ChatGPT and other tools in naturalistic writing contexts. This ecological perspective helped reveal that ChatGPT-mediated writing in non-interventionist and self-initiated contexts involves a range of software tools and multiple behaviors associated with each tool's use.

The two-level coding scheme—tools and behaviors—constructed from observations and stimulated recall of participants' writing sessions illustrates the complex and dynamic nature of ChatGPT-mediated writing. At the tool level, participants employed various digital tools alongside ChatGPT, guided by individual preferences and writing community norms. As noted in L2 writing studies in naturalistic contexts (Séror and Gentil 2023, Yoon 2016), L2 writers often rely on multiple resources to meet evolving writing needs. This suggests that rather than viewing ChatGPT as a standalone tool for L2 writing (Barrot 2023, Su et al. 2023), it may be more ecologically valid to examine ChatGPT's use in conjunction with other tools and to consider how these relationships affect its affordances and constraints in L2 writing processes. At the behavioral level, participants demonstrated that using ChatGPT in L2 writing involved not only diverse gaze and non-gaze behaviors but also individualized sequences of these behaviors to meet specific writing goals. More specifically, ChatGPT-related behaviors, such as *ChatGPT prompting*, *ChatGPT output reading*, and *incorporating ChatGPT output into writing*, indicate that ChatGPT can reshape conventional L2 writing processes by introducing additional procedural steps.

The analysis of the two participants, Collin and Thomas, using timescapes and frequency/duration tables of their writing processes, reveals idiosyncratic patterns in tool usage and behavior. Collin's ChatGPT use was brief and targeted; he tended to copy and paste segments of his original work without detailed prompts and engaged minimally with ChatGPT's output. Instead, he favored re-reading and revising independently in Overleaf. This restrained use of ChatGPT reflects his view of ChatGPT as a reference tool for vocabulary and expressions rather than for content generation or paraphrasing. His relatively high English proficiency may shape this approach, as he could maintain an independent revision process and exercise greater control over his work with minimal reliance on AI tools. In contrast, Thomas's writing sessions demonstrate how ChatGPT-related behaviors shifted according to the purpose and stage of the writing process. For brainstorming and outlining, he briefly scanned ChatGPT's outputs; however, when elaborating on the outline he drafted, he reviewed outputs more carefully to expand his outlines. Notably, Thomas used a Text Document Comparison Site along with a zig-zagging gaze pattern to compare changes between his original text and ChatGPT's suggestions. His gaze movement between ChatGPT and either Overleaf or the Text Document Comparison Site on a split screen enabled him to incorporate revisions thoughtfully. This zig-zagging pattern suggests that Thomas was willing to regulate his use of ChatGPT for his writing goals rather than automatically accepting its suggestions for revision.

The study also illustrates that these idiosyncratic tool use patterns and behaviors are not only affected by individual writing goals and perceptions of tools, including ChatGPT, but also conditioned by the arrangement of digital software. Collin's brief and copy-and-paste-focused ChatGPT interactions were facilitated by his separate tab setup for ChatGPT and Overleaf, which he felt helped him concentrate on revising a final draft. In contrast, Thomas's side-by-side screen setup supported sustained engagement and detailed prompting throughout brainstorming, planning, and drafting. This observation highlights the diverse ways in which L2 writers utilize online tools at various stages of the writing process (Li 2012) and the strategic organization of digital tools to scaffold their writing more effectively (Hort 2020, Overstreet 2022, Séror and Gentil 2023).

Along with its implications for ChatGPT-mediated writing in naturalistic settings, this study offers methodological insights. Webcam-based eye-tracking facilitated observation of L2 writers' naturalistic engagement with ChatGPT. This study shows that using the webcam, an easily accessible tool on participants' personal devices, webcam-based eye-tracking can collect data in authentic and everyday settings and enhance the ecological validity of the research. This study finds webcam-based eye-tracking particularly effective for observing digital practices, as it captures both visual attention to the screen and physical actions like typing and mouse control and produces gaze plot recordings as data (Godfroid et al. 2020). Furthermore, the study used gaze plot recordings and stimulated recalls to create timescapes visualizing L2 writers' real-time evolution of micro-writing processes (Smith 2017). This visualization identified general trends and sequential behaviors in digital writing processes, enabling comparative analysis of participants' patterns. Despite its benefits, timescape visualization has limitations in displaying very short-duration behaviors and precise durations and frequencies of individual codes. To address these limitations, frequency/duration tables were thus used as a complementary data presentation method. As writing processes in digital spaces grow increasingly complex and dynamic, there is a pressing need to develop innovative data analysis and presentation methods.

As a qualitative study exploring L2 writers' naturalistic use of ChatGPT, this study provides several pedagogical implications. The findings suggest that ChatGPT, as a general-purpose tool, encourages diverse and emergent behaviors as it flexibly adapts to various contexts and needs. This adaptability, however, can bring a level of unpredictability that can complicate structured educational applications. Specifically, this unpredictability may pose challenges to developing a step-by-step framework for integrating ChatGPT into L2 writing instruction. Even though there are increasing efforts to regulate the ways of using AI tools such as ChatGPT in educational contexts, this study argues that, given ChatGPT's ubiquity in everyday life, it is necessary to explore diverse ChatGPT usage patterns both in and beyond the classroom for effective and sustainable integration of AI tools in L2 writing education.

6. Conclusion

Using webcam-based eye-tracking as a qualitative data collection method for a naturalistic language learning experience alongside other conventional qualitative research methods, this study investigated South Korean STEM graduate students' use of ChatGPT in self-initiated writing. By unobtrusively tracing L2 writers' real-time evolutions during the writing process, this study suggests that ChatGPT may urge us to rethink conventional writing stages and behaviors (Godwin-Jones 2022, Warschauer et al. 2023). Given that new writing or reference tools have always affected perceptions and practices of writing (Haas 1996), further empirical studies could explore how ChatGPT or other AI-based tools are used in relation to other tools or materials and how such relational perspectives on the use of tools shed light on the changing practices of L2 writing.

References

Bailey, C. and J. Withers. 2018. What can screen capture reveal about students' use of software tools when undertaking a paraphrasing task? *Journal of Academic Writing* 8(2), 176-190.

Barrot, J. S. 2023. Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing* 57, 100745.

Chan, C. K. Y. and W. Hu. 2023. Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education* 20(1), 43.

Gánem-Gutiérrez, G. A. and A. Gilmore. 2018. Tracking the real-time evolution of a writing event: Second language writers at different proficiency levels. *Language Learning* 68(2), 469-506.

Godfroid, A. 2019. *Eye Tracking in Second Language Acquisition and Bilingualism: A Research Synthesis and Methodological Guide*. Routledge.

Godfroid, A., P. Winke and K. Conklin. 2020. Exploring the depths of second language processing with eye tracking: An introduction. *Second Language Research* 36(3), 243-255.

Godwin-Jones, R. 2022. Partnering with AI: Intelligent writing assistance and instructed language learning. *Language Learning & Technology* 26(2), 5-24.

Hamel, M. -J., J. Séror and C. Dion. 2015. *Writers in Action: Modelling and Scaffolding Second Language Learners' Writing Process.* Higher Education Quality Council of Ontario.

Haas, C. 1996. *Writing Technology: Studies on the Materiality of Literacy*. Routledge.

Hayes, J. R. 1996. A new framework for understanding cognition and affect in writing. In C.M. Levy and S. Ransdell, eds., *The Science of Writing: Theories, Methods, Individual Differences and Applications* (pp. 1-27). Routledge.

Hellmich, E. A. and K. Vinall. 2023. Student use and instructor beliefs: Machine translation in language education. *Language Learning & Technology* 27*,* 1-27.

Hort, S. 2020. Digital writing, word processors and operations in texts: How student writers use digital resources in academic writing processes. *Journal of Academic Writing* 10(1), 43-58.

Kessler, M. 2024. *Digital Multimodal Composing: Connecting Theory, Research and Practice in Second Language Acquisition*. Multilingual Matters.

Kim, N. 2020. A study on the current state of academic English journal writing by graduate engineering students. *Cross-Cultural Studies* 59, 55-80.

Lafford, B. A. 2009. Toward an ecological CALL: Update to Garrett (1991). *Modern Language Journal 93*(S1), 673-696.

Latif, M. M. M. A. 2019. Eye-tracking in recent L2 learner process research: A review of areas, issues, and methodological approaches. *System* 83, 25-35.

Leijten, M. and L. Van Waes. 2013. Keystroke logging in writing research: Using Inputlog to analyze and visualize writing processes. *Written Communication* 30(3), 358-392.

Leijten, M., L. V. Waes, K. Schriver and J. R. Hayes. 2014. Writing in the workplace: Constructing documents using multiple digital sources. *Journal of Writing Research* 5(3), 285-337.

Li, Y. 2012. Undergraduate students searching and reading Web sources for writing. *Educational Medial Interanational* 49(3), 201-215.

Liu, G. and C. Ma. 2024. Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model. *Innovation in Language Learning and Teaching* 18(2), 125-138.

Liu, S. and G. Yu. 2022. L2 learners' engagement with automated feedback: An eye-tracking study. *Language Learning & Technology* 26(2), 78-105.

Manchón, R. M. and J. Roca De Larios. 2023. *Research Methods in the Study of L2 Writing Processes* (Vol. 5). John Benjamins Publishing Company.

Michel, M., A. Révész, X. Lu, N. -E. Kourtali, M. Lee and L. Borges. 2020. Investigating L2 writing processes across independent and integrated tasks: A mixed-methods study. *Second Language Research* 36(3), 307-334.

Overstreet, M. 2022. Writing as extended mind: Recentering cognition, rethinking tool use. *Computers and Compositions* 63, Article 102700.

Révész, A., M. Michel and M. Lee. 2019. Exploring second language writers' pausing and revision behaviors: A mixed-methods study. *Studies in Second Language Acquisition* 41(3), 605-631.

Roca de Larios, J., R. Manchón, L. Murphy and J. Marín. 2008. The foreign language writer's strategic behaviour in the allocation of time to writing processes. *Journal of Second Language Writing* 17(1), 30-47.

Sauro, S. and K. Zourou. 2019. What are the digital wilds? *Language Learning & Technology* 23(1), 1-7.

Séror, J. and G. Gentil. 2023. Direct observation of writing activity: Screen capture technologies. In R. M. Manchón and J. Roca de Larios, eds, *Research Methods in the Study of L2 Writing Processes,* 161-182. John Benjamins.

Smith, B. E. 2017. Composing across modes: A comparative analysis of adolescents' multimodal composing processes. *Learning, Media and Technology* 42(3), 259-278.

Stickler, U. and L. Shi. 2017. Eyetracking methodology in SCMC: A tool for empowering learning and teaching. *ReCALL* 29(2), 1601-77.

Strauss, A. and J. Corbin. 1998. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* (2nd ed.). Sage Publications, Inc.

Su, Y., Y. Lin and C. Lai. 2023. Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing* 57, 100752.

Takayoshi, P. 2016. Methodological challenges to researching composing processes in a new literacy context. *Literacy in Composition Studies* 4(1), 1-23.

van Weijen, D., H. van den Bergh, G. Rijlaarsdam. and T. Sanders. 2009. L1 use during L2 writing: An empirical study of a complex phenomenon. *Journal of Second Language Writing* 18(4), 235-250.

Warschauer, M., W. Tseng, S. Yim, T. Webster, S. Jacob, Q. Du and T. Tate. 2023. The affordances and contradictions of AI-generated text for writers of English as a second or foreign language. *Journal of Second Language Writing* 62, 101071.

Yan, D. 2023. Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies* 28(11), 13943-13967.

Yang, X. and I. Krajbich. 2021. Webcam-based online eye-tracking for behavioral research. *Judgment and Decision Making* 16(6), 1485-1505.

Yoon, C. 2016. Individual differences in online reference resource consultation: Case studies of Korean ESL graduate writers. *Journal of Second Language Writing* 32, 67-80.

Zou, M. and I. Huang. 2023. The impact of ChatGPT on L2 writing and expected responses: Voice from doctoral students. *Education and Information Technologies*.

Examples in: English
Applicable Languages: English
Applicable Level: All

Appendix A. The Coding Scheme for Tools in ChatGPT-Mediated Writing

Category	Code	Description
Tools	1. ChatGPT	A generative AI chatbot developed by OpenAI.
	2. Overleaf (https://ko.overleaf.com/)	An online-based collaborative LaTeX editor extensively used within scientific communities for writing, editing, and publishing documents. Automatically formats writer's manuscripts in the style required for various journals. All participants in this study voluntarily utilized Overleaf.
	3. Google	A search engine that offers a wide range of services. Participants mainly used it to search for words, gather information, and as a gateway to other websites.
	4. Dynalist (https://dynalist.io/)	An online-based outlining tool designed for creating structured lists and organizing ideas. Its hierarchical structure is particularly useful for compiling complex information, making it ideal for brainstorming.
	5. Text Document Comparison Site (https://wepplication.github.io/tools/compareDoc/)	A web application that compares two documents and highlights the differences between them using colors, enabling users to intuitively identify discrepancies between two versions of a document.
	6. ArXivGPT (https://chromewebstore.google.com/detail/solar-arxiv/fbbfpcjhnnklhmncjickdipdlhoddjoh?pli=1)	A ChatGPT extension for the ArXiv website. Summarizes the paper uploaded in ArXiv according to the prompts that the user set up through ChatGPT.
	7. Notion	A software application that serves as an all-in-one workspace for documentation, data management, task management, and project planning. Users can create customized pages that may include text, database tables, images, and videos.
	8. Others: Naver Dictionary, PDF	Tools consulted with less frequency (once or twice), with less relevance to ChatGPT usage.

Appendix B. The Coding Scheme for Behaviors in ChatGPT-Mediated Writing

Category	Code	Code Description
Gaze Behavior	1. ChatGPT output scanning	Fixate the words at the beginning of lines and swiftly move the gaze horizontally.
	2. ChatGPT output reading	Fixate on all or most words presented by ChatGPT.
	3. Zig-zagging	Continuously shift gaze back and forth between two software arranged on a split screen.
	4. Re-reading	Re-read segments of the text produced by oneself.
	5. Overleaf compile and reading or scanning	Compile the document in Overleaf and read or scan the preview.
	6. Software scanning	Scan the content in software other than ChatGPT. No typing or reading.
Non-Gaze Behavior	7. ChatGPT prompting	Type texts into the input field of ChatGPT placed at the bottom of the chat window.
	8. Text construction (K)	Type actual words in Korean
	9. Text construction (E)	Type actual words in English
	10. Revising writing	Revise texts written by oneself at word, sentence, or paragraph level.
	11. Drag and drop	Copy text from one software (excluding ChatGPT) and paste it into another (including ChatGPT), such as from Notion to ChatGPT or Overleaf.
	12. Incorporating ChatGPT output into writing	Transfer text from ChatGPT to another software by copying and pasting or typing manually.
	13. Revising ChatGPT output	Revise ChatGPT output after copying and pasting it into another software.
	14. Search	Search for information on Google.
	15. Locating	Change the spatial organization and align the software.
	16. Grammarly suggestion apply	Click on the Grammarly extension embedded in Chrome.
	17. Other	Activities not related to writing.