JournalAIde: Empowering Older Adults in Digital Journal Writing

Shixu Zhou
The Hong Kong University of Science and Technology (Guangzhou)
Guangzhou, China
Hong Kong University of Science and Technology
Hong Kong, China
szhouav@connect.ust.hk

Weiyue Lin
Faculty of Education
The University of Hong Kong
Hong Kong S.A.R., China
weiyuelin@connect.hku.hk

Zuyu Xu
Hong Kong University of Science and Technology
Hong Kong, China
zxubo@connect.ust.hk

Xiaoying Wei
Hong Kong University of Science and Technology
IIP(Computational Media and Arts)
Hong Kong, China
xweias@connect.ust.hk

Raoyi Huang
WInE
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA
raoyih@andrew.cmu.edu

Xiaojuan Ma
Hong Kong University of Science and Technology
Hong Kong, Hong Kong
mxj@cse.ust.hk

Mingming Fan*
The Hong Kong University of Science and Technology (Guangzhou)
Guangzhou, China
The Hong Kong University of Science and Technology
Hong Kong, China
mingmingfan@ust.hk

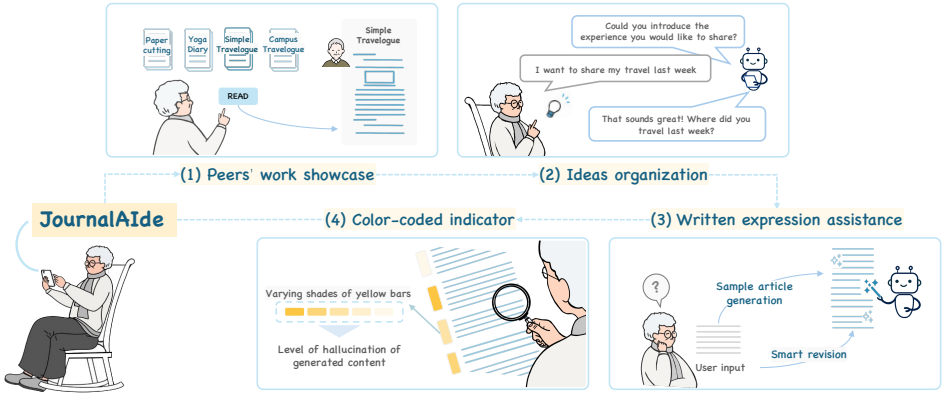


Figure 1: Overview of JournalAIde system. It supports older adults in digital journaling by: (1) Boosting their confidence by showcasing the writing output of their peers who have similar writing backgrounds using JournalAIde as vicarious experience. (2) Helping organize ideas and sustain attention with a question and answering chatbot. (3) Facilitating written articulation with LLM-powered text generation. (4) Providing visual cues to help them edit the generated text and tailor it to their preference.

*Corresponding author

CHI '25, Yokohama, Japan

Abstract

Digital journaling offers a means for older adults to express themselves, document their lives, and engage in self-reflection, contributing to the maintenance of cognitive function and social connectivity. Although previous works have investigated the motivations

ACM ISBN 979-8-4007-1394-1/25/04
https://doi.org/10.1145/3706598.3713339

and benefits of digital journaling for older adults, little technical support has been designed to offer assistance. We conducted a formative study with older adults and uncovered their encountered challenges and preferences for technical support. Informed by the findings, we designed a Large Language Model (LLM) empowered tool, JournalAIde, which provides vicarious experience, idea organization, sample text generation, and visual editing cues to enhance older adults' confidence, writing ability, and sustained attention during digital journaling. Through a between-subjects study and a field deployment, we demonstrated the JournalAIde's significant effectiveness compared to a baseline system in empowering older adults in digital journaling. We further investigated older adults' experiences and perceptions of LLM writing assistance.

CCS Concepts

• **Human-centered computing** → **Interactive systems and tools**; **Empirical studies in HCI**; **Accessibility technologies**.

Keywords

Older adults' creativity; Digital journal writing; Accessibility

ACM Reference Format:

Shixu Zhou, Weiyue Lin, Zuyu Xu, Xiaoying Wei, Raoyi Huang, Xiaojuan Ma, and Mingming Fan. 2025. JournalAIde: Empowering Older Adults in Digital Journal Writing. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26–May 01, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 19 pages. https://doi.org/10.1145/3706598.3713339

1 Introduction

Many older adults are increasingly taking active roles in creating digital content, from videos and podcasts to digital journals including travel blogs, expressive writings and self-reflective journals [7, 27, 28, 51, 76]. As digital content creators, older adults could enjoy many benefits, including self-expression, enhanced digital skills, maintenance of social connections, and increased digital participation [78, 88]. Particularly, *digital journaling* in various topics and formats has become a popular way for older adults to document their lives, share opinions and views, and release repressed emotions [7, 64, 83]. Moreover, older adults often approach journaling with a desire to preserve memories and wisdom for future generations [65]. This practice not only strengthens their connections with family and community but also fosters intergenerational engagement [46, 57], as shared reflections and values spark conversations and build bridges between generations [67]. Digital journaling also offers meaningful engagement in later life and has proven instrumental for older adults to develop new identities, acquire social interaction, and stay cognitively active [7, 16, 61].

Despite the benefits, digital journaling is a challenging task, which involves deep reflection and creativity in writing processes as well as writing skills [2, 28, 83]. Within the broader context of facilitating the general writing process, previous studies have introduced various methods such as word embedding, transformer architecture, and language models for text understanding and generation [10, 53, 86]. Existing works have also explored writing support systems based on the algorithms and investigated their performance and effects. For instance, Writing Mentor [12], an application that can evaluate writing performance automatically, was developed to improve students' writing skills. Likewise, MepsBot [71] incorporates assessment and recommendation modes to help caregivers create text comments on online mental health communities. More recently, WordCraft [93] was designed to aid creative story writing in collaborative scenarios. However, most of these works were aimed at young students and professional writers in different fields, leaving out older adults in their design and evaluation processes. For older adults, during the journaling process, they need tools designed to support their natural changes in cognition and memory, which often shape their different writing performance and pace [40, 42, 63, 94]. Furthermore, for the topics of digital journaling, older adults tend to prefer content centered around their life experiences and share within their community [61, 83], which is a more personal and free writing context that is different from the professional and educational settings. Such differences may also introduce unique needs that require more tailored designs.

While some prior work proposed ways to facilitate digital content creation among older adults, little has focused on supporting digital journal writing. For instance, workshops and courses have been conducted to foster video production and appropriation skills among older adults [27]. Enmesh was developed to facilitate new relationship building among older adults through photos and messages creation and sharing [88]. Although those works are informative, they fell short of providing support for digital journal writing, which requires more introspection and writing skills of older adults to express themselves. Regarding support related to digital journal writing, the xPress platform [8] enables older adults with decreased vision to engage in online blogging communities through voice input. InMyDays [26] offers a digital diary tool to encourage older adults' self-care. However, these works either focused exclusively on specific subcategories like support for older adults with vision loss or provided platforms limited to basic chronological archives of daily emotions and events, lacking creative writing assistance for older adults. The process of digital journaling often involves more uses of multimedia content and complex creative writing processes, including multiple rounds of planning, translating, and reviewing [29, 82], which makes the existing works less applicable to the digital journaling context. Thus, there remains a significant gap in understanding the practical challenges older adults may encounter in digital journaling and in designing tools that are specifically tailored to alleviate those challenges.

To fill the research gap, we conducted a formative study with older adults who had experience in writing digital journals to understand the challenges they encountered and possible technical support during the process. Using semi-structured interviews and demonstrations of generative AI technologies, we engaged participants in exploring how technical support might better address their needs. The findings revealed that older adults prefer technical support that provides personalized assistance, helping them rather than taking over entirely. And the challenges uncovered mainly fall into three areas: lack of confidence, deficiency in writing ability to organize and articulate ideas, and difficulty in staying engaged for sustained attention. Based on these findings, we designed and implemented JournalAIde, an interactive system empowered by large language model (LLM) to support older adults in digital journaling by **1)** Boosting their confidence by showcasing the writing output of their peers who have similar writing backgrounds using

JournalAIde as vicarious experience, **2)** Helping organize ideas and sustain attention with a question and answering chatbot, **3)** Facilitating written articulation with LLM-powered text generation, **4)** Providing visual cues to help them edit the generated text and tailor it to their preference.

Finally, we first conducted a between-subjects user study to evaluate the effectiveness of JournalAIde compared to a Baseline system in supporting older adults' digital journaling. We recruited 24 older adults to create a digital journal using the assigned system. Results showed that JournalAIde significantly enhanced older adults' confidence and writing ability, fostered a more engaging experience compared to the Baseline, as well as maintained good system usability. Following the between-subjects user study, we conducted a field deployment study with 6 older adults to investigate the user experiences and perception of the LLM assistance in a real-world setting. We then further discussed insights and design considerations for future systems like JournalAIde.

The contributions of this work are as follows: **1)** JournalAIde, an interactive system designed to assist older adults in digital journaling. **2)** A comprehensive between-subjects study that demonstrates the effectiveness of JournalAIde compared with a Baseline system. **3)** A field deployment study exploring user experiences with JournalAIde and perception of LLM writing assistance in a real-world setting. **4)** Insights and design considerations that can guide the future design of interactive systems aimed at supporting digital journaling for older adults.

2 Related Work

Our work is informed by the digital content creation and digital journaling among older adults and existing supporting tools for writing and digital journaling within the broader context.

2.1 Digital content creation and digital journaling for older adults

An increasing number of older adults have participated in digital content creation nowadays under the influence of multiple factors, such as the advance of digital technology and the popularity of social media [11]. This involvement can range from blogging [7, 15, 16], creating videos [27, 34, 81], to producing podcasts [76, 77]. When older adults engage in digital content creation, they often bring a wealth of life experiences and unique perspectives that can enrich online communities [79]. In addition, previous works have revealed multifarious benefits for older adults who partake in digital content creation, such as fostering self-expression [15, 88], improving digital skills [84, 97], and maintaining social connections [32, 87]. To investigate the experiences and perceived benefits of older adults who created digital stories, Hausknecht *et al.* monitored older adults creating and sharing digital stories over a ten-week course and identified benefits including social bonding, life retrospection, and legacy building. The findings suggested that digital storytelling enhances emotional and social well-being among older adults [36]. Participating in digital content creation such as blogging can also give older adults a larger platform and community in which to network with audiences with similar interests and create new relationships [52]. Considering the rich benefits, increasing HCI research has paid attention to the support for older adults' digital content creation. For example, Keisari *et al.* conducted an online creative process of digital photocollage for 24 Italian and Israeli community-dwelling older adults aged 78 to 92 based on creative arts therapies [47]. The findings showed that digital photos and photocollages could help older adults emotionally engage in storytelling, safely interact with therapists, and reflect on their lives, which further supports crucial late-life development. In the project "Older Voice" [76], Reuter *et al.* identified challenges and introduced digital solutions to help older adults in the UK engage more effectively in media production, which includes enhancing audience interaction through digital means and introducing a digital production tool, demonstrating that customized digital processes can facilitate older adults' meaningful participation in community media.

Among various types of digital content creation, digital journaling including travel blogs, self-reflective and expressive writings is a particularly beneficial activity for older adults, combining the advanced features of digital technology with the benefits of traditional journal writing [7, 28, 51]. Journaling can become a tool for life review, enabling them to reflect on experiences that may have shaped their identity, beliefs, and values [3, 74]. Brady and Sky examined the history, current habits, and benefits of journaling among 15 older learners, revealing keeping journaling helps older individuals cope with daily life, experience the joy of discovery, and foster personal voice and spirit [61]. Differing from those of younger age groups, the process of reflection in journaling for older adults is more tied to a desire to preserve memories and wisdom in a form that can later be shared with others, especially with younger generations [65]. Additionally, older adults' journaling is more relational and intertwined with family and community [46]. It can strengthen their relationships with families and friends, which can spark conversations and encourage social engagement [57]. This relational aspect of journaling also serves as a bridge between generations, as older adults record memories or values they wish to pass down [67]. During journaling, the process of reflection and creative writing can also help keep older adults' brains active, which is beneficial for preserving cognitive function. Campbell studied the effects of journaling for 8 weeks on well older adults who exhibited symptoms of depression [13]. The results showed that journaling can be an effective form of cognitive therapy to help older adults reduce depression through memory and reflection. Engaging in journaling promotes mental agility and provides a structured outlet for older adults to organize their thoughts, revisit memories, and document personal insights [61]. This form of writing not only encourages cognitive exercise but also helps build confidence by empowering individuals to articulate their experiences in meaningful ways [83].

Despite the benefits, journaling as a creative activity can be challenging for older adults because it requires reflective thinking, self-expression, and the ability to organize thoughts into a coherent narrative [39]. Additionally, older adults may struggle with organizing thoughts coherently or holding onto an idea long enough to record it [35, 45]. Although research has shown that older adults have unique motivations to keep journaling and can greatly benefit from it, limited support is provided in prior research to address the challenges that may arise for them when attempting to participate in this activity. To empower older adults to engage more effectively in digital journaling, in this work, we aim to design an interactive

tool that supports digital journaling specifically tailored to their needs.

2.2 Support tools for writing and digital journaling

Journal is a type of writing, which is often reflective and used to record thoughts, feelings, and observations over time [37]. Previous research has proposed a number of tools that focus on different needs for general writing support, including some LLM-powered tools in enhancing various user's writing practices. DiaryMate [50] is an LLM-driven journaling assistant that supports users in introspective writing by automatically generating relevant sentences based on the existing text or user-provided keywords. Another LLM-powered journaling app, MindfulDiary [49], was developed to assist psychiatric patients in documenting daily experiences through conversational interactions, designed with input from mental health professionals. Yuan *et al.* introduced Wordcraft [93], a collaborative text editor integrating a large neural language model to enhance the story writing process. They found that their system facilitated unique co-writing experiences by engaging users in dialogue, fulfilling specific stylistic requests, and offering creative suggestions to help users who get blocked. Fan *et al.* collaborated with older adults and observed how they modified existing instructions and subsequently summarized guidelines for writing senior-friendly instructions [25]. Peng *et al.* presented MepsBot [71], a writing assistant for online mental health community supporters, featuring assessment and recommendation modes. Another system named Writing Mentor [12] utilized NLP techniques to offer automated writing evaluation feedback to enhance students' writing skills in postsecondary settings. It addresses various writing aspects like source use, coherence, and English conventions and collects data on students' writing self-efficacy through an optional survey.

Although there is a growing body of research on writing support, most has focused on the general population or specific context, with limited support for older adults' journal writing. Previous research suggests that older adults may have unique needs for tools tailored support that acknowledges natural shifts in cognition, vision, and motor abilities with age [35, 45, 60]. These shifts can influence their writing approach, such as preferring more deliberate organization or adapting to a slower pace that allows for thoughtful reflection [40]. Moreover, considering the unique motivations of older adults for journaling, as discussed previously, it is crucial to provide tools that bolster their confidence and encourage consistent journaling practices. Nevertheless, there is limited research focused on developing support tools to assist older adults with digital journaling. Fernandez *et al.* developed InMyDay [26] digital diary tool specifically designed for older adults, with the aim of facilitating self-care and self-reflection by enabling them to document their daily activities and emotions. However, it only arranges daily entries in a predefined format and allows users to record in chronological order, which limits the free and creative thinking and organization of digital journaling. To address the challenges faced by low-vison older adults, Brewer and Piper developed xPress [8], an accessible and voice-based blog community specifically tailored for older adults with acquired vision loss. Nevertheless, this work only focuses on the needs of older adults with low vision, which is a subcategory of the overall needs of older adults.

To sum up, there is a discernible research gap regarding the challenges older adults face during the digital journaling process and the technical support they expect. In this work, we utilized a formative study to uncover the challenges faced by older adults and their preferences regarding technical support. Subsequently, we developed JournalAIde, an LLM-empowered interactive system for supporting digital journaling for older adults.

3 Formative Study

3.1 Participants and Procedure

With the ethical approval of our institution, we recruited 9 participants (six female, three male; age range 55-70; all retired; indexed P1 to P9) for the formative study through online advertising, social media, and word-of-mouth at a local senior community. All participants had at least a junior high school degree and reported experience writing digital journals while being novice journal writers (fewer than 3 writings in the past 6 months). We first conducted semi-structured interviews with participants to learn about their experience in writing digital journals. More specifically, we asked about their perceptions of digital journals and the challenges they encountered during such a process. Furthermore, to explore potential technical support for digital journaling among older adults, we introduced participants to various generative AI tools related to digital journaling. In particular, we showcased the text generation using ChatGPT[1] (with model GPT-3.5) for their chosen topics and provided them with the opportunities to experiment with the image generation utilizing MidJourney[2]. Subsequently, we discussed how those tools could be leveraged to facilitate digital journaling practices. Each session took about 40-60 minutes, and all participants received compensation according to local standards.

The interview sessions were audio-recorded and transcribed into text. We employed thematic analysis [6], following an inductive coding approach to generate initial codes from the data. Two researchers first independently coded the transcripts, and then the research team refined and organized the codes following the framework of writing processes (e.g., planning, drafting, revising) as identified in [29, 82]. Disagreements were resolved through discussion during weekly meetings. Data saturation was reached after the seventh interview, as no new codes emerged in the final two interviews. Therefore, we stopped to conduct more interviews. Four key themes were identified about older adults' encountered challenges and preference for technical support during digital journaling. Specifically, the themes include "lack of confidence", "deficiency in writing ability", "difficulty maintaining sustained attention", and the preference for "collaborative assistance preserving personal writing styles".

3.2 Findings

Generally speaking, all participants agreed that taking digital journals is a good way to document and share their lives and reflections, as well as cognitive training to keep their minds sharp. Participants used digital journals to record various events in their lives, including but not limited to traveling experiences, reunion activities with

[1] https://chat.openai.com/
[2] https://www.midjourney.com/

old friends, and the growth of offspring. Based on the transcript of the discussion, we identified the challenges in three categories for older adults during their digital journal writing process: lack of confidence, deficiency of writing ability in organizing and articulating ideas, as well as difficulty in staying engaged for sustained attention. In addition, we distinguished older adults' desire for technical support that provides personalized assistance, helping them rather than taking over entirely. The details of our formative study findings are shown in Table 1.

3.2.1 Lack of confidence. Many participants (5/9) indicated that sometimes they felt their writing was not good enough compared to those highly read digital journals shared in some online senior communities, which hindered them from writing their own digital journals (**C1**). *"I will read other people's travel articles on the online senior communities, but they are often very well written and literary. I feel like I can never write as well as them."* P1 said. Participants also mentioned their concerns from perfectionism and uncertainty about their digital journals. Specifically, as older adults are often willing to share their journals with their community and the next generations, they want to bring out the best of themselves in digital journals. Failing to meet their own high standards and requirements during writing or get enough attention after sharing would often lead older adults to doubt their ability and the value of their writing (**C2**), thereby affecting their confidence in digital journal writing. For instance, P3 said, *"I would like to present the best of my life, and I would keep revising and editing it until I am satisfied."*

3.2.2 Deficiency of writing ability. All participants mentioned that they experienced challenges with their ability to write digital journals. Those challenges regarding their writing ability can be further divided into two categories: organizing the writing logic and articulating their thoughts in written words. More specifically, most participants (8/9) found it challenging to begin the writing process due to difficulty organizing their ideas and establishing a logical structure for expression (**C3**). Such problems may be especially pronounced in older adults, who often face age-related declines like working memory, making it harder for them to organize their thoughts coherently [40, 69]. *"Sometimes there are life moments that I want to record by writing digital journals, but I don't know how to get started and what to write about first."* P7 said. Additionally, many participants (7/9) also noted that it was difficult to find the appropriate words and construct sentences to express their ideas effectively in writing (**C4**). *"I tried to write digital journals to record my life, but I found it hard to write beautiful words to describe it effectively."* P2 said. Consequently, participants may view their writings as too mundane and ultimately give up. As previous research has noted, older adults in particular may experience difficulties with word retrieval and sentence construction due to age-related cognitive changes [33, 40].

3.2.3 Difficulty to stay engaged for sustained attention. According to many participants (7/9), staying engaged during digital journaling is identified as a significant challenge. Specifically, participants mentioned encountering difficulties midway through the writing and not knowing how to further develop their content. Such problems may lead to frustration, loss of focus and interest, becoming more pessimistic about the outcome of the writing, and a reluctance to continue (**C5**). P6 explicitly mentioned that *"I need to sit there for a long time to write. My mind would wander easily and I feel like giving up when I encounter some difficulties during writing."* Furthermore, the difficulty is compounded by the solitary process of digital journaling. Some participants (4/9) reported the absence of collaboration as an additional challenge that makes writing a more lonely process (**C6**). For older adults, the extended solitary task can feel particularly demanding, especially without timely external motivation and support. This observation also aligns with previous research which highlights the importance of collaboration on older adults' engagement in other forms of digital content creation [78].

3.2.4 Technical Support Tailored to Personal Needs. After exploring the fundamental capabilities of generative AI tools in both image and text generation, participants found text generation features to be powerful tools supporting their digital journaling. However, they expressed that image generation was not as useful, given that they predominantly use photos taken by themselves in digital journals. Nevertheless, many participants (6/9) also expressed their concerns about the over-reliance on the AI text generation feature and the potential impact on their self-autonomy (**C7**). *"I want to reflect and recall the moments during my writing. AI should assist me rather than take over all the tasks."* P4 said. This preference may be related to older adults viewing journaling as a cognitive training activity that requires active mental engagement and emphasizes authorship, which is often discussed in the context of human-AI collaboration [16, 38]. Another specific need for technical support is about the AI-generated text. Five participants explicitly mentioned that the AI-generated text is "too polished" and contains imaginary parts, which is a kind of hallucination often detected in large language model-generated content [95]. Given this deviation from their personal writing styles, participants also emphasized the necessity for technical support to better accommodate their personal preferences and help preserve their writing styles (**C8**).

4 JournalAIde System

Based on the findings from Formative Study and existing research, we derived the following 4 design goals (**DG**s) to further guide the design of our JournalAIde system.

DG1: Provide vicarious experience and boost older adults' confidence to write. Previous studies have identified instances of writing apprehension within students [1, 21]. Similarly, participants in our formative study generally set high standards for their digital journals (**C1**) and would doubt their writing value and become unconfident about their writing ability if they fail to meet the standards (**C2**). Consequently, they would feel hesitant and anxious to get started with digital journaling. Following self-efficacy theory [4], JournalAIde system aims to boost older adults' confidence by providing vicarious experience that on the one hand reduces their writing apprehension and on the other hand demonstrates the system's ability to encourage them to write [70].

DG2: Modularize the writing task and provide help at every stage. Writing is a complex task that involves many processes and is an essential step in digital journaling [29]. Previous research has found that due to the decline of working memory, it is more difficult for older adults to handle

Table 1: Identified older adults' challenges and needs for technical support in digital journaling

Categories	Challenges/Needs
Confidence	**C1**. Get overwhelmed and discouraged by highly read writings
	C2. Doubt their writing ability and writing value
Ability	**C3**. Organize ideas and establish writing logic
	C4. Articulate ideas with effective words
Engagement	**C5**. Sustain attention and interest when faced with writing difficulties
	C6. Engaging in the writing process without collaboration
Technical Support	**C7**. Assist users rather than take over all tasks
	C8. Accommodate personal preferences and help preserve writing styles

both writing structure and coherence at the same time [40]. Furthermore, participants in our formative study also indicated facing difficulties with organizing and articulating ideas (**C3, C4**). Therefore, we aim to modularize the writing process into planning, articulating, and revising processes, and provide appropriate assistance at each writing stage.

DG3: Incorporate AI to facilitate collaboration and instill positive outcome expectations. In human-computer interaction, user engagement has always been valued and important [23]. Previous work has also shown that increasing user engagement can make users more willing to complete the task and perform better in various writing tasks [56, 96]. Participants in our formative study also mentioned their challenges of losing attention and feeling alone during their digital journaling. By incorporating AI like chatbot and smart assistant to facilitate collaboration and instill positive expectations, user engagement can be enhanced in terms of control, interest, and interactivity [56, 66], which allows users to become more engaged (**C5, C6**) and confident in the digital journaling process (**C2**) [22].

DG4: Support sample text generation and intuitive revising on the generated content. Text generation has been used as a support feature for various writing-related tasks and has proven useful and powerful [17, 93]. Based on our formative study, participants also recognized the potential of generative AI tools like ChatGPT in their digital journaling and expressed expectations and concerns about the agency and hallucination. In response to those insights, JournalAIde system is designed to provide support for sample article generation and intuitive revision of generated content. Such support can enable users to engage in co-authoring experiences with generative AI tools (**C7**) and help users articulate their ideas (**C4**) and edit them to personal preference and style (**C8**).

4.1 User Interface

Based on the design goals, we developed the JournalAIde system (Figure 2, 3; Interfaces translated into English, original interfaces in the Appendix B), incorporating three views: Square View, Edit View, and Chatbot View, which correspond to the three tabs in the system (Figure 2, ⓪): Article Plaza, Edit, and Smart Assistant, respectively. Older adults can freely switch and use the three views (tabs) during their digital journaling. Next, we introduce the main features.

4.1.1 Square View. Square View (Figure 2, (A)) provides a place where older adults can view the writing output of peers with similar writing backgrounds using JournalAIde. Specifically, we demonstrated the benefits of digital journaling for older adults using text prompts and presented four randomly selected journals from previous participants' work to provide vicarious experiences and encourage older adults to write (**DG1**). To ensure that all older adults participating in the user study could read four journals from their peers, we included the works of older participants in our pilot study (N=4) into the random sampling pool. Users could click the corresponding button on the journal card to read the journal (Figure 2, ①) and can directly click "Start Writing" (Figure 2, ②) to navigate to the Edit View.

4.1.2 Edit View. For the Edit View (Figure 2, (C) and Figure 3, (C)), we divided the writing content into separate paragraph blocks with each block being an image or a piece of text, and provided basic editing functions. Users can create a new paragraph to get started and insert text or images between any blocks by clicking the plus sign (Figure 2, ③), and they can also adjust the order by simply dragging the blocks. After the user inserts an image, a caption is automatically generated under the image in the small gray font (Figure 2, ④) with the GPT-4 Vision[3] feature. If users get stuck during the writing process, they can click the "Chat with Smart Assistant" button (Figure 2, ⑤) to navigate to the Chatbot View for help in idea organization and articulation.

In addition, users can choose smart editing (Figure 3, ⑥) powered by GPT-4 Chat Completion[4] feature when revising a paragraph (all prompts used in JournalAIde are shown in the Appendix A), and modify and polish the paragraph by inputting their revision ideas (Figure 3, (B1)). To further assist users in revising the generated text (**DG4**), we integrated a color-coded indicator (Figure 3, ⑦) for the generated content. Different shades of yellow are used to indicate the consistency between the generated text and the user's input, with lighter shades indicating greater coherence and darker shades indicating greater divergence. The specific details of the color-coded indicator algorithm are illustrated in 4.1.4.

4.1.3 Chatbot View. We implemented the Chatbot View (Figure 3, (A)) with integrated chatbot and sample text generation capabilities

[3]https://platform.openai.com/docs/guides/vision
[4]https://platform.openai.com/docs/guides/text-generation/chat-completions-api

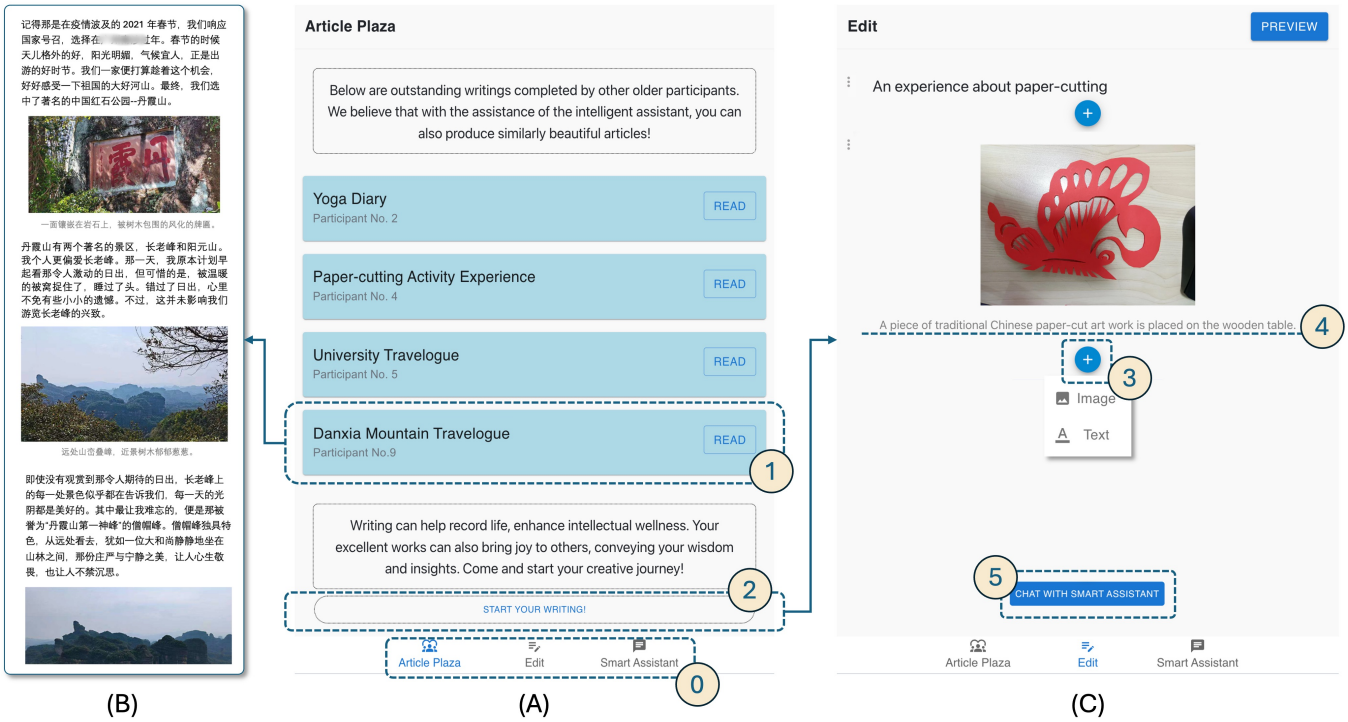


Figure 2: (A) User interface (UI) of Square View. ⓪Older adults could freely explore and switch these three tabs, ①view the writing outputs by their peers, and ②navigate to Edit View to start writing. (B) Example of participant's writing output. (C) Part of UI of Edit View. Older adults could ③add text and image blocks and ④ caption would be automatically generated after a new image is inserted. ⑤When users get stuck, they can navigate to Chatbot View for assistance.

to foster users' digital journaling, as chatbots have been proven to increase user confidence, provide engaging experiences and facilitate content creation [14, 56, 90] (**DG2**, **DG3**). The chatbot is designed to guide users in reflecting and describing the content they want to write about through conversation and questioning (e.g., recalling a memorable experience or expressing opinions on a specific matter). We leveraged the GPT-4 Chat Completion API to develop this chatbot, with a tailored prompt to ask questions about the content described by older adults.

At any time during interaction with the chatbot, user can choose to generate sample text (Figure 3, ⑧) based on their existing chat history (**DG4**). For the generated sample text (Figure 3, (B2)), users can either enter their thoughts for a full-text regeneration or merge the generated text into the Edit View as a reference or for further revisions (Figure 3, ⑨). The generated text would be split based on paragraphs and appended into the Edit View as blocks.

4.1.4 Color-coded Indicator. We utilized varying shades of yellow bar assigned to each paragraph to indicate the consistency between generated text and user input information. Following prior research on hallucination detection in LLMs [41, 58], We assess consistency by inputting the generated text and user input information into the GPT-4 model for a new round of inference and scoring. To validate the strategies and metrics used, we further prompted the GPT-4 model and obtained evaluation metrics including content and detail consistency, emotional and sentiment consistency, and contextual and situational consistency. These metrics align with those identified in prior works [43, 75, 91]. The logic flow of the color-coded indicator implementation is shown in Figure 4, and Table 2 specifies user input information data under different cases. Specifically, each generated paragraph would be segmented into sentences, and each sentence would be compared to user input information individually. The GPT-4 is prompted to give a decimal consistency score between 0 and 1 for each sentence, and the final score of each paragraph is the weighted average of the scores obtained for each sentence. With the final consistency score, we visualize the color-indicator by adjusting the saturation value in HSV (hue, saturation, value) color system while keeping the hue and lightness constants ($hue = 43°, saturation = 1 - consistencyScore, value = 0.98$). In addition, if the user modifies a paragraph him/herself, the consistency score of the paragraph will be set to 1. And if a paragraph already has a consistency score, the final consistency score after a prompt-based modification (Smart editing in Edit View and Full-text regeneration in Chatbot View) will be set to the average of the old score and the newly calculated score.

5 Study 1: Between-subjects Study in Lab Setting

To evaluate the JournalAIde system, we conducted a between-subjects study (N=24) to understand its effectiveness in supporting older adults' digital journaling compared with a Baseline system.

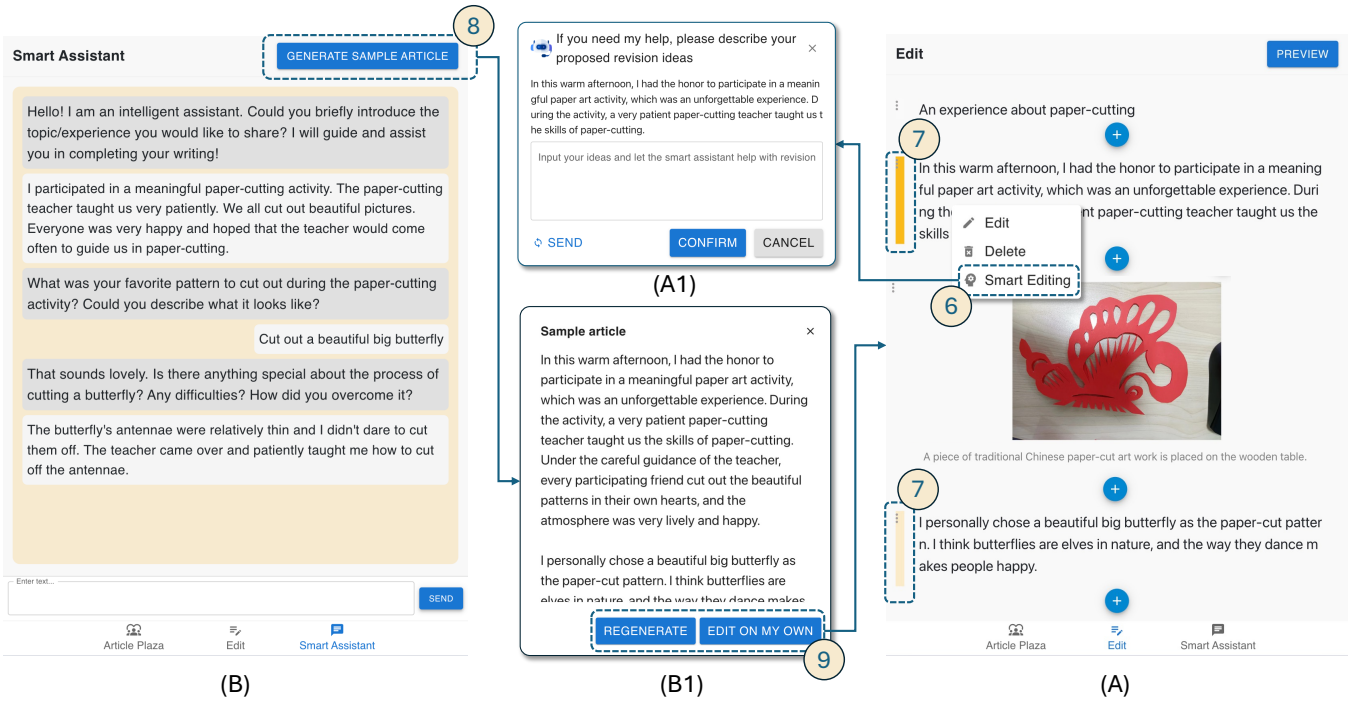


Figure 3: (A, A1) Part of UI of Edit View. Older adults could ⑥choose "smart editing" to trigger prompt-based paragraph revision assistance and ⑦refer to the color-coded indicators that reflect consistency between user input information and generated content for editing (to be better aware of and control generative AI's hallucination.) (B, B1) UI of Chatbot View. The chatbot would guide older adults in idea organization. At any time, users could ask for ⑧sample text generation help and ⑨request a full-text regeneration with their revision ideas or add the generated content to Edit View.

Table 2: Different user input information cases for consistency score calculating

View	Cases	User Input Information
Chatbot View	Sample Text Generation	Conversation History
	Full-text Regeneration	Conversation History + Revision Idea
Edit View	Paragraph Smart Editing	Original Paragraph + Revision Idea

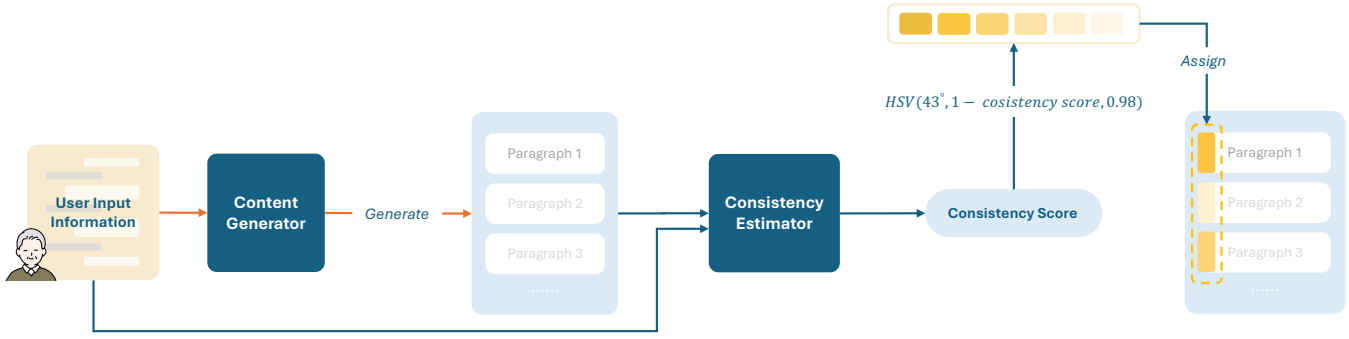


Figure 4: Logic flow of the color-coded indicator. JournalAIde assesses consistency between user input and generated text using GPT-4 scores for each sentence, and the paragraph's score is a weighted average of those. Different shades of yellow are used to indicate the score.

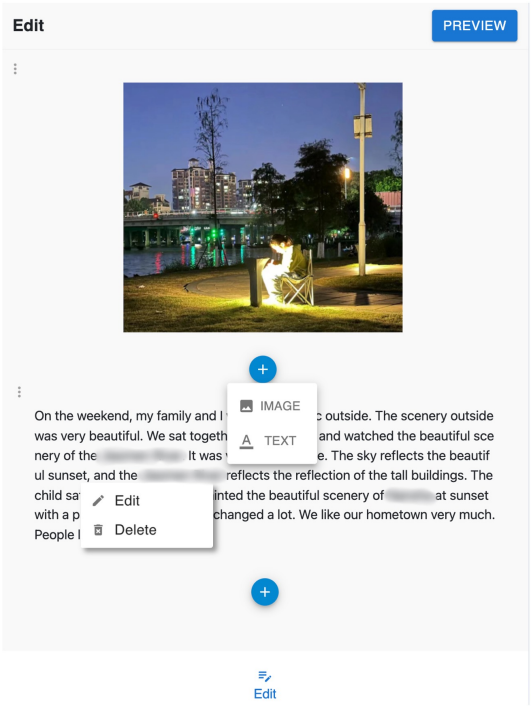


Figure 5: UI of the Baseline system. Older adults can insert images, add/edit text blocks, and arrange elements' order via drag and drop.

In the following sections, we will discuss our design of the Baseline system, participant recruitment, study procedure, as well as evaluation metrics and hypotheses.

5.1 Baseline System

To understand the effectiveness of JournalAIde and fully explore the incorporation of generative AI technology in older adults' digital journaling, we compared JournalAIde with a Baseline system that only offers basic Edit View to preserve basic editing capabilities. Specifically, in the Baseline system we removed all LLM-based functionalities and color indicators, retaining basic functionality such as text and image input, modification, and the ability to drag and rearrange blocks of content. Users can use these reserved features for creating digital journals.

5.2 Participants Recruitment & Study Procedure

We recruited 24 older participants (20 female, 4 male; age range 60-84) through online advertising, word-of-mouth, and snowball sampling at a local senior community. To mitigate the potential impact of variations in individual writing abilities and experiences, participants were randomly assigned to experimental (indexed P1 to P12) and baseline groups, with 12 participants in each group, to interact with different systems. During the user study, participants were first introduced to the basic features and usage of the system, followed by a practice and a question-and-answer session. After participants became familiar with how to use the system, they began creating digital journals using the assigned system. After completing the writing of digital journal, participants were asked to fill out a questionnaire about their interaction experiences. They were subsequently interviewed to elaborate on their questionnaire responses, provide feedback on the system, and express their expectations for future system improvements.

5.3 Evaluation Metrics and Hypotheses

Drawing from the evaluation methodologies and frameworks utilized in prior studies [44, 55, 89], we assessed the systems based on System Usefulness encompassing User Empowerment and Engagement and Writing Outcome Analysis, as well as System Usability. Details of the evaluation metrics and hypotheses are as follows:

5.3.1 User Empowerment and Engagement. In response to the insights drawn from the formative study and related work, we assessed the systems' assistance in boosting confidence, enhancing writing ability, and fostering engagement throughout the digital journaling process. Specifically, we inquired participants about their perceived overall writing ability support and assistance in organizing ideas and choosing words for writing as a measure of writing ability support. To measure user engagement, we adopted Brien's engagement model [68] as a theoretical framework and derived the questionnaire with reference to previous work [56, 92].

5.3.2 Writing Outcome Analysis. Considering digital journaling as a predominantly subjective task involving sharing and recording life experiences [61], we conducted an outcome analysis by asking users to subjectively rate their writing output on overall quality, accuracy in expressing emotions and thoughts, and willingness to share it with others.

5.3.3 System Usability. In interactive support systems, the addition of helpful features often leads to system complexity and difficulty in learning and using [31]. To assess system usability, we followed previous work and adopted a shorter version of the System Usability Scale [9, 56] to evaluate three aspects: ease of use, ease of learning, and willingness to reuse.

5.3.4 Hypotheses. JournalAIde is expected to boost the confidence and writing ability of older adults while providing a more engaging experience during digital journaling and maintain good system usability. Therefore, we formulated the following hypotheses:

- **H1** Compared to the baseline, JournalAIde more effectively supports older adults' confidence (*H1a*), ability (*H1b*), and foster engagement (*H1c*) during the digital journaling process.
- **H2** Compared to the baseline, JournalAIde demonstrates a more significant improvement in older adults' digital journal writing outcomes in terms of self-perceived overall quality (*H2a*), accuracy of expression (*H2b*), and willingness to share (*H2c*).
- **H3** Compared to the baseline, JournalAIde exhibits better system usability regarding "ease of use" (*H3a*), "ease of learning" (*H3b*) and "willingness to reuse" (*H3c*).

6 Results of Study 1

We collected participants' ratings on the support for their confidence, writing ability, and experience engagement, as well as the self-perceived quality of their writing outcomes during the digital

journaling process. Additionally, we also gathered participants' ratings on the system usability during the interaction with JournalAIde and Baseline systems, as well as participants' subjective ratings on their writing outputs. All the questions in the questionnaire follow the 7-point Likert scale, where 1 represents the most negative stance (strongly disagree) and 7 represents the most positive stance (strongly agree).

We conducted the Mann-Whitney U test [59] on participants' different ratings on the JournalAIde and Baseline systems to gain a comprehensive understanding and assessment. Following previous work [18, 30, 85], we calculated the test statistic, p-value, and the corresponding effect size *r* for each hypothesis, as shown in Table 3. Under the current sample size (N=24), we used G*Power[5] to conduct post hoc power analysis. All hypotheses in H1 were accepted with power level greater than 0.85 and other accepted hypotheses reached power level of at least 0.81. Therefore, the sample size we selected was adequate for our evaluation.

6.1 System Usefulness

6.1.1 User Empowerment and Engagement. As shown in Figure 6, in comparison to the Baseline system, the Mann-Whitney U test scores indicate that participants felt significantly more supported in terms of their confidence ($U = 141.00, p < 0.001$) and ability ($U = 144.00, p < 0.001$) when using JournalAIde for digital journaling. Therefore, *H1a* and *H1b* are accepted. Furthermore, participants using JournalAIde also reported significantly higher levels of engagement ($U = 130.50, p < 0.001$) as illustrated in Figure 7a, *H1c* accepted. According to Cohen's proposed effect size *r* and the threshold [18], the effect sizes of the test on confidence (*Eff. Size* = 0.83), ability (*Eff. Size* = 0.86), and engagement (*Eff. Size* = 0.69) all show a large effect (>= 0.5), which implies the significant improvement magnitude about user empowerment and engagement of JournalAIde.

6.1.2 Writing Outcome Analysis. Regarding the writing outcome, participants using JournalAIde reported significantly higher levels of self-perceived overall quality ($U = 136.00, p < 0.001$), accuracy of expression ($U = 120.50, p = 0.002$) and willingness to share ($U = 134.00, p < 0.001$) compared to those using the Baseline system (Figure 7b), leading to the full acceptance of **H2**. In general, most participants (9/12) explicitly mentioned that JournalAIde helped improve their overall writing quality and made them more willing to document and share their lives in the digital journaling form.

6.1.3 User Feedback on Functional Views. To better understand the roles and effects of different views (Square, Edit, and Chatbot View) and their features in system usefulness, we also asked participants in interviews to elaborate on their evaluation of the specific views and how they perceived the support provided.

Square View. Most participants (9/12) indicated that the Square View inspired their interest in starting a digital journal and fostered their willingness to learn from peers. Specifically, participants commented that viewing peers' work with comparable levels could *"boost their confidence"* and even *"spark competitive thoughts"* to begin writing. Participants also expressed a desire to *"be coaxed into writing"* (P2, P3, P12) after reading others' entries and to have a more global assistant with features like a cute avatar and voice interactions to *"invite participation."* When discussing journal sharing and writing communities, participants mentioned their preference for customizable sharing options, such as *"family and friends"* (P2, P4) and *"interest groups like reading and poetry writing"* (P12). Additionally, some also expressed interest in collaborative writing, hoping to *"work with friends to learn and improve their writing"* (P5, P11).

Edit View. All participants (12/12) chose to first generate some sample text using the Chatbot View and then modify it in the Edit View. There were totally 56 paragraphs generated by the LLM of which 27 paragraphs were edited by participants. Additionally, more than half of the participants (7/12) expressed that the color-coded indicator could serve as a reference for them to locate the "AI mistakes" and revise paragraphs involving AI-generated content. To gain a deeper understanding of how participants interacted with color-coded indicators, we tracked the consistency scores for generated content as well as participants' editing activities. The results show that participants tend to edit paragraphs with relatively lower consistency scores, as shown in Figure 8. Notably, the average consistency score for generated paragraphs was 0.71, and participants made 21 edits to paragraphs with consistency scores below this average and only 6 edits to content above this average. Such results also reflected participants using color-coded indicators as a reference when editing AI-generated content.

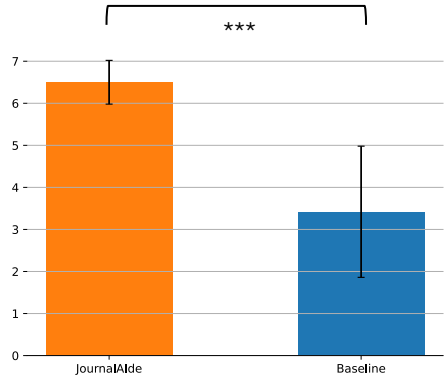
Furthermore, although participants expressed a desire for AI to further help revise their writings during interviews, only a minority of participants (3/12) used the smart editing feature. Participants mentioned that despite having watched the example use cases, they were still unsure about *"how to formulate the prompt"* (P11) and *"how AI can help"* (P10), which hindered them from using such prompt-based regeneration features. Regarding the automatic photo captioning, participants generally found it a useful feature, but they also noted that the generated captions needed further refinement as they sometimes failed to consider the context, resulting in cold descriptions that were purely informative and lacked emotional expression. For example, P2 said, *"I would prefer to write 'Aerial yoga, my favorite' than 'A woman is doing a yoga pose'."*

Chatbot View. Regarding the Chatbot View, most participants (10/12) expressed that the question-and-answer sharing mode with the chatbot helped them to recall details, guided their thinking direction, and the sample text generation also helped their writing ability in language expression. P8 said, *"It can guide the direction of my thinking and memories and provide some beautiful and rich expressions"*. However, participants also mentioned cases where the chatbot's questioning was either *"divergent from the expected writing direction"* or *"went too much into details"*. For example, when P4 compared food from her travel destination to her hometown, the chatbot asked about her perception of the latter. *"It is different from my expected writing topic [emphasis on travel experiences]."*, explained P4. Also, there were also cases where the LLM did not preserve the user's emotions. For instance, P9 described an unpleasant experience, but the generated text downplayed this negative emotion and sublimated it into an unpleasant but also memorable experience, which could be attributed to the restriction policy of existing LLM services to avoid the generation of negative content.

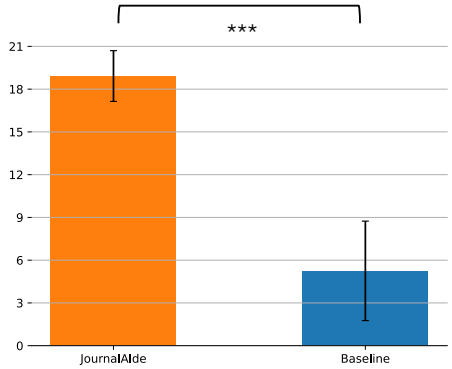
[5]https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower

Table 3: The statistical analysis of user feedback using Baseline and JournalAIde, where the p-value (+: .050 < p < .100, *:p < .050, **:p < .010, ***:p < .001) and effect size are reported.

Category	Factor	Baseline Mean(S.D.)	JournalAIde Mean(S.D.)	Statistics				Hypothesis
				U	p-value	Sig.	Eff. Size	
Experience	Confidence	3.42(1.56)	6.50(0.52)	141.00	<0.001	***	0.83	H1a accept
	Ability	5.25(3.49)	18.92(1.78)	144.00	<0.001	***	0.86	H1b accept
	Engagement	24.33(7.64)	37.17(3.33)	130.50	<0.001	***	0.69	H1c accept
Outcome	Overall quality	3.42(1.24)	5.92(0.67)	136.00	<0.001	***	0.77	H2a accept
	Accuracy of expression	4.33(1.61)	6.33(0.65)	120.50	0.002	**	0.58	H2b accept
	Willingness to share	3.17(1.27)	6.08(1.08)	134.00	<0.001	***	0.75	H2c accept
Usability	Ease of use	5.50(1.00)	6.17(1.03)	99.00	0.056	+	0.33	H3a reject
	Ease of learning	5.67(0.89)	6.00(1.48)	94.50	0.092	+	0.27	H3b reject
	Willingness to reuse	2.92(1.83)	6.33(0.65)	135.50	<0.001	***	0.76	H3c accept

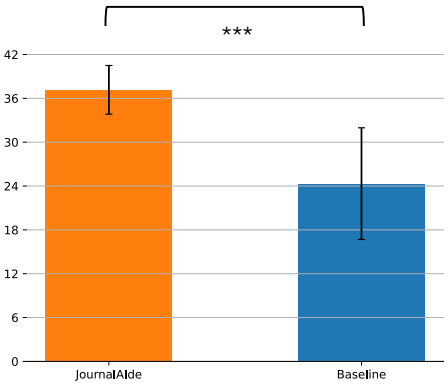


(a) Ratings on confidence support

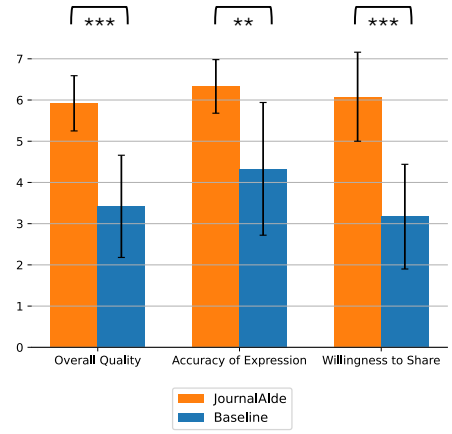


(b) Ratings on writing ability support

Figure 6: Participants' ratings of confidence and writing ability support of JournalAIde and Baseline (*: $p < .001$)**



(a) Ratings on experience engagement



(b) Ratings on writing output

Figure 7: Participants' ratings of their experience engagement and writing output (: $p < .010$, ***: $p < .001$)**

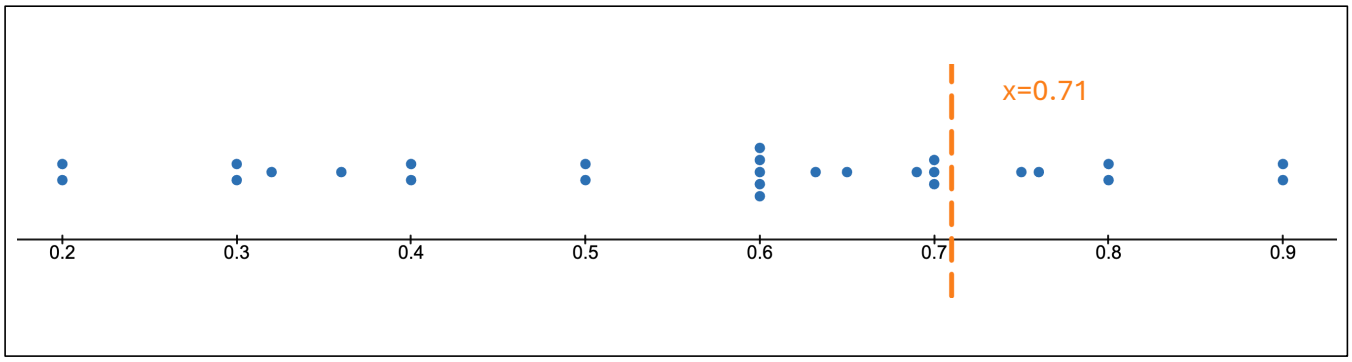


Figure 8: Participants' editing behavior on AI-generated content. The horizontal axis represents the consistency score between a participant's input and AI's generated content (see section 4.1.4 for calculation details), while the dots signify editing operations.

Additionally, participants reported a preference for AI-generated content to align with their personal style, while having an acceptable "polishing threshold"(P2, P7, P8). P2 explained, *"I appreciate elegant expression, but if my writing is at a fifth-grade level, then polishing it to a junior high school level is enough, I don't want it to be too exaggerated."* In terms of engagement, four participants (P2, P7-8, P12) explicitly mentioned that seeing the articles generated increased their confidence and enjoyment in writing. *"It gives me a sense of control over the writing process, which makes me more confident"*, added P8.

6.2 System Usability

Participants expressed a significantly higher willingness to use JournalAIde again ($U = 135.50, p < 0.001$), *H3c* accepted. However, the rating of "ease of use" is only marginally higher ($U = 99.00, p = 0.056$), and there is no significant improvement regarding "ease of learning" ($U = 94.50, p = 0.092$) when comparing JournalAIde to the baseline. Therefore, *H3a* is marginally accepted and *H3b* is rejected. Despite this, JournalAIde still received high average scores in terms of "ease of use" ($Mean = 6.17, SD = 1.03$) and "ease of learning" ($Mean = 6.00, SD = 1.48$), indicating its good usability.

Moreover, additional features in JournalAIde introduce more learning costs during the participants' initial attempts to use it, especially for the commonly less tech-savvy older adult population. *"This tool is useful, but I need time to learn and become familiar with its functions to use it effectively"*, said P2. Four participants also explicitly expressed their unfamiliarity and uncertainty with AI-related functions (P1, P7, P10, P11), which may impose more learning effort. P1 mentioned, *"I rarely use AI tools in my daily life, so it will take me some time to become familiar with the usage patterns before I can use it with confidence."*

In summary, JournalAIde demonstrated significant effectiveness in empowering older adults' digital journaling regarding confidence, writing ability, and engagement in our lab setting user study. Participants appreciated the Chatbot View's question-and-answering sharing and the sample text generation features, which helped them organize and articulate their ideas more effectively. They also noted that viewing peers' work and collaborating with the chatbot assistant increased their confidence and made the writing process more enjoyable. However, participants also indicated a need for more time to familiarize themselves with the system and AI functionalities. Building on prior research that highlight the value of field studies [62], we are motivated to further explore the effect and usage of JournalAIde in more natural environments through a field deployment study.

7 Study 2: Field Deployment Study

To better understand the effects and usage patterns of JournalAIde, we conducted a second study in the form of a 10-day field deployment with 6 older adult participants.

7.1 Updates on JournalAIde System

To enhance the field deployment study while preserving the overall design, we made several improvements to JournalAIde. First, we implemented basic login functionality, user management features, and logging to track users' interactions with different features. We also modified Square View to display writings by our field study participants. Additionally, results from Study 1 showed that older adults had difficulty with the prompt-based regeneration features and found image captions disconnected from the context. To address these issues, we analyzed commonly used revision types in writing and tailored them to our context [24], designing prompt templates for our regeneration features. The templates were sampled from three revision categories: (1) formal changes, (2) meaning-preserving changes, and (3) meaning changes. We upgraded the LLM to GPT-4o[6], leveraging its text and vision capabilities to improve image captioning by incorporating both the image content and the user's existing text.

7.2 Participants and Study Procedure

We held two informational sessions for individuals interested in the study, which was advertised online, through word-of-mouth in a local senior community, and via follow-up invitations to participants from Study 1 who had expressed further interest. As a result, six participants (5 female, 1 male; age range 60–78; 2 from Study 1; indexed P1 to P6) formally signed up for the study. We introduced how to use the JournalAIde system and encouraged participants to complete at least four writings using it. To better assist older

[6]https://openai.com/index/hello-gpt-4o/

adults' use of our system, we also distributed the image text rich handout and made a WeChat group for any questions and discussions. After completing the 10-day study, participants were invited to participate in interviews. Drawing insights from previous work on human-AI collaboration [38, 50], older adults' content creation, and results of our Study 1, our interview protocol focused on: **1)** Participants' existing journaling habits **2)** Perceived support for user empowerment when using the system, **(3)** Impressions and user experience with JournalAIde, **(4)** Perceptions of agency, authorship, and reliance in collaborating with LLM, and **(5)** Expected features and improvements for the system.

8 Results of Study 2

Totally, we have 36 pieces of writings composed by participants, except one participant who completed 3 writings, all other members finished at least 4 writings.

8.1 User Empowerment: Confidence, Motivation and Writing Ability

All participants indicated that our JournalAIde system supports their confidence and writing abilities. In terms of confidence and motivation, participants noted that the system boosts their confidence in trying *"different writing genres"* (P5). And it also encourages them to *"get motivated and curious about learning new technologies"* (P2) introduced through writing activities, where their interests in both writing and technology mutually reinforce each other. Furthermore, participants highlighted their diverse goals for sharing of writings. These include using sharing to motivate themselves to persist in their learning activities (like calligraphy, painting, and chorus, commonly popular in participants), documenting positive life experiences to inspire others, fostering community connections through shared topics, and updating distant relatives on recent events. Our design of Square View can also facilitate and expand such writing and sharing activities and the associated benefits. For instance, P5 shared that part of her motivation for writing and sharing is learning from other participants. *"Seeing others write today makes me feel more motivated to write as well."*, P5 added.

Regarding writing ability, participants expressed that our system can *"improve efficiency"* (P2-3), *"help polish expressions"* (P1, P6) and *"elevate the theme of the writing"* (P5). Nevertheless, when it comes to *"deep psychological thoughts"* and *"nuanced descriptions"*, participants found that the system's effectiveness in capturing and describing is relatively limited (P2-4). Such perceived limitation could be partly attributed to the current information collection process, which primarily relies on the question-and-answering in the Chatbot View. As participants sometimes only provide short responses or keywords, it may lead to insufficient capture. Additionally, participants highlighted the need for tailored support to various writing genres (P2, P5). For instance, different topics like "documenting travel experiences" and "expressing personal reflections" require different structures and thematic focuses. For long-term development of writing ability, participants acknowledged that while AI can assist their writing, genuine improvement in writing ability requires consistent practice, time and accumulation of experience.

8.2 Older Adults' Interaction with LLM

Overall, our participants engaged in a total of 576 messages with the LLM in the Chatbot View and used the regeneration feature 24 times.

Chatbot Q&A. All participants agreed that the chatbot's Q&A feature effectively guided their thinking and helped expand their ideas and enrich their perspectives when describing events. It also encouraged them to think and reflect (P1-2). However, as observed in Study 1, the chatbot occasionally asked questions that some participants felt diverged from their intended focus. This issue was discussed further during the interviews. In terms of the question quality, participants still noted that there were cases where the questions lacked specificity and direction (P4), leaving them confused and uncertain about the quality of the subsequent generated text (P1). Additionally, although the Chatbot View was designed to allow participants to generate sample text at any point during the conversation, some participants reported that the chatbot consistently asked questions from different perspectives. This often prompted them to respond to the questions instinctively (P2-3, P6), which might affect their perception of the question quality, especially when older adults have already shared a lot, but the questions continue to come up.

Prompt-based Regeneration. For the regeneration assistance, participants still reported a low use rate and most of them relied on the provided prompt templates rather than crafting their own (3/24). Some participants expressed uncertainty about how to use the feature effectively (P1-2, P6) and a fear of making mistakes (P2). Specifically, for the prompt templates, the type of "Meaning-preserving Changes" was leveraged most, whereas the other two types were used far less often. Details on the number of uses and specific prompt template examples can be found in Table 4.

Color-coded Indicator. Regarding the color-coded indicator, participants reported either not noticing it at all or noticing it but not using it as a reference. This contrasts with the results of Study 1, which may be attributed to the lab setting where we provided help and explanations about the features midway through to help keep their memory of the features fresh. And older adults may be less sensitive to subtle design cues like this. Although the indicator was not used much, participants (P3, P5) still highlighted their sensitivity to errors in journals and expressed a desire for a feature that could highlight areas where the LLM is uncertain and may contain knowledge errors.

Image Captioning. In terms of the updated image caption, all participants praised it for accurately capturing the information in the image and producing beautiful, contextual descriptions. Inspired by this, participants (P1, P4) expressed interest in a future feature that allows them to upload photos first and then engage in conversations with the chatbot. Such a feature could help users recall memories while also providing the LLM with a clearer contextual understanding.

Usage Patterns, Behaviors, and Existing Journaling Habits. From the interviews and usage logs, we observed different usage patterns between participants with a "Regular Journaling Habit" (P2, P5) and those engaging in "Event-Driven Journaling" (other participants). Participants with event-driven journaling habits primarily used the Chatbot View to initiate new journal entries. On the

Table 4: Revision Types, Number of Uses, and Prompt Template Examples

Revision Types	# of Uses	Template Examples
Formal Changes	1	Please check and correct my typos and punctuation errors.
	2	Please check and fix my grammatical issues.
Meaning-preserving Changes	6	Help me remove unnecessary words to make this writing more concise while maintaining its meaning.
	10	Help me rewrite this passage to make it more engaging while preserving its original meaning.
Meaning Changes	2	Help me add some vivid descriptions to expand this piece of writing.
	0	Add some details to make the background of this writing clearer.

other hand, participants with regular journaling habits would also import their previous writings into the system, seeking assistance to refine and polish their content, as well as utilize features like image captioning to enrich their existing entries. These differences extended to their expectations for LLM assistance. Participants with regular journaling habit expected feature that provides writing frameworks or helps refine their initial drafts. Nevertheless, those with event-driven journaling habits leaned toward the sample text generation, where the LLM could create drafts based on their input, which they would later revise and modify to suit their needs.

8.3 Agency, Authorship, and Reliance When Collaborating with LLM

Although participants initially reported instinctively answering the Chatbot's questions, some gradually developed proactive control over the collaboration with the LLM. From the chat logs, we observed instances where participants attempted to "change the topic" (P4) or instructed the chatbot to "skip this part" (P1). Specifically, P4 added, *"It's important to have a rough outline beforehand, and actively provide key points and information [to the AI] to guide the content and control the direction of writing."*

Regarding authorship, participants expressed conflicting opinions. While some attributed authorship primarily to themselves (P1-4), others felt the AI's major contribution made the work feel less like their own (P5-6). Additionally, P3 took a more moderate view, stating that there was an expectation of writing and *"If the AI introduces significant new content beyond what I originally imagined, it feels like it doesn't belong to me. If it's consistent with my expectations, then it still feels like it belongs to me, even with a lot of polish."*

Similarly, although all participants agreed that writing served both to document daily life and as a form of cognitive training, their perspectives on reliance on AI and its impact on cognitive training varied. Some participants expressed little concern about cognitive training being diminished during writing, arguing that AI primarily helps reduce the burden of corrections and descriptions, while the thinking process is still their own (P4). They indicated that they would actively revise and create content themselves (P3-4) while also learning from the AI (P1). However, some participants also expressed concerns that excessive reliance on AI could lead to cognitive decline, making them mentally lazy over time (P2, P5-6). In addition, these different perspectives may also be related to the older adults' existing journaling habits, position of and purpose for writing. As P2 mentioned, some people may sometimes just want to *"write a few sentences and let the AI generate an article"*, while others prefer to *"fully engage in the process and write more content themselves"*.

9 Discussion

Previous work on older adults' journal writing has primarily demonstrated the benefits of writing through workshops [51, 57] or focused on specific support subcategories, such as voice-based assistance for those with vision loss, or basic daily activity archives [8, 26], lacking creative writing assistance. Our work first identified the challenges that older adults face in the digital journaling process, along with their preferences for technical support. Then we developed and explored a support approach that encompasses vicarious experience-based confidence building and LLM powered writing assistance. Through user studies in both lab and field deployment settings, we demonstrated the effectiveness and significance of JournalAIde while also identifying several key discussion topics, such as agency and transparency of AI writing assistance for older adults, individual differences like existing journaling habits and positioning of journaling in support requirements, and opportunities for collaboration and sharing among peers. We discuss the insights and design considerations derived from our user studies findings, as well as the limitations and future work.

9.1 Agency and Transparency of AI Writing Assistance for Older Adults

Our results suggest that older adults may instinctively continue answering the LLM chatbot's questions, where some of the unexpected questions could get them confused. This tendency could be related to older adults' unfamiliarity with the technology [5, 19] and preference for thorough reading and learning over skimming the content [48]. Specifically, the unfamiliarity with the technology may make it challenging for them to understand the reasoning process behind the chatbot, while their preference not to skim content could make them hesitant to jump out from the conversation. In this case, even if the LLM chatbot may eventually collect sufficient information after multiple rounds of conversation, such incidents may cause confusion, affecting older adults' engagement and interaction with AI assistance [68]. To alleviate this issue, strategies can be designed from different directions: Enhancing the specificity of the chatbot's questions and Increasing the transparency of the system perceived by older adults.

Enhancing the Specificity of Chatbot Questions: Drawing insights from our study findings, future systems could allow older adults to pre-input images and describe their ideas along with them. This is an approach they can start using the LLM chatbot aligning with their existing habits and preferences for creating image-centric social media posts [73]. Furthermore, this approach could not only enhance LLM's understanding of contextual information to improve question and text generation, but also mirror a familiar interaction flow and intuitive user experience for older adults, thereby contributing to their engagement during interactions.

Increasing System Transparency for Older Adults: To increase older adults' perceived transparency and reduce confusion when encountering unexpected chatbot questions, future systems could investigate ways to provide a writing framework and visualize the writing progress. For example, the system could implement a flowchart to envision the writing path, which allows users to preview the existing writing plan and several possible future directions and steps, enabling them to make their preferred choices. Additionally, the follow-up questions from the chatbot could also be improved by offering alternative categories and options, giving users more control over the interaction. In this way, the system could help older adults understand the writing trajectory and reduce the feelings of uncertainty and getting lost.

9.2 Explicit and Automated Assistance for Older Adults

In our JournalAIde system, we implemented the color-coded indicator to help older adults adjust the generated content to their preferences, and prompt-based regeneration to support idea input and content revision. Although the color-coded indicator was found to be useful in Study 1, likely due to the controlled lab setting, both features were reported to be either unnoticed or infrequently used during our field deployment. Specifically, the limited use of the color-coded indicator may be attributed to the system already requiring older adults to adapt to and learn unfamiliar functionalities. It might be difficult for older adults to take such subtle visualization into account at the same time, especially during a cognitively demanding task like writing [40, 83]. On the other hand, the low usage of the regenerative feature may result from the high degree of freedom it provides. As participants mentioned, they were hesitant to try it out of fear of making mistakes or uncertainty about the results. This is further evidenced by their preference to use the provided prompt templates rather than create prompts independently in our field deployment.

Future systems could offer more intuitive and explicit cues for identifying potential mistakes made by the LLM. For example, as participants suggested, the color-coded indicator could highlight words or sentences that might contain knowledge errors or hallucinations. Tapping on these highlighted sections could then reveal the potential issues along with suggested directions for revision. For regeneration, future work could explore more contextualized prompt templates that adapt to the older adults' input. Additionally, after the user writes a section, the system could proactively offer lightweight suggestions, such as word and phrase adjustments, coherence improvements, or clarity checks. This approach would reduce the needs on user-initiated prompts, making the feature more seamless and accessible for older adults.

9.3 Older adults' Individual Factors and Comfort Level for Writing Assistance

Our study results suggest that older adults' existing journaling habits and their positioning of the writing task may influence their comfort levels with technical assistance, as well as their perceptions of authorship and cognitive reliance. For instance, participants with regular journaling habits in our field deployment expressed concerns that long-term use of such tools might lead to reduced cognitive effort, contradicting their intention to use digital journaling as a form of cognitive training. In contrast, others viewed the tools as valuable for efficiently capturing life events in written form. Participants also expressed varied preferences regarding the desired features of LLM-based assistance.

To address these differences and better accommodate with older adults' diverse preferences and comfort levels, future work could offer tailored assistance levels based on individuals' goals and writing habits. For instance, future systems could collect older adults' information about existing journaling habits, literacy levels, previous writings, and their expectations for the current journaling experience. For those prioritizing efficiency, stronger assistance could provide more automation of generation, revision, and polish of text. For those focused on cognitive engagement, lighter assistance may involve question scaffolding and providing writing frameworks, rather than direct generating texts to guide deeper exploration and learning, as well as proposing new topic for writing reflection. Additionally, incorporating such user-specific information would enable LLM to better understand individual writing styles and provide more personalized support, enhancing writing empowerment and satisfaction.

9.4 Engagement and Motivation for Older Adults to Participate in Digital Journaling

Previous research has found that older adults are often motivated to adopt digital journaling through the influence and encouragement of others [15]. To enhance older adults' confidence and motivate them to engage in digital journaling, JournalAIde offered the Square View to communicate the benefits of digital journaling in text form and showcase the works of other participants according to the self-efficacy and expectancy theories for writing [54, 70]. In our studies, participants indicated that viewing other participants' writings was interesting and motivated them to start writing themselves while also mentioned the desire to collaborate with friends to learn from and support each other. Therefore, future work could include more collaborative and social support features, such as providing common topics for group members to reflect on and write together, as well as exchanging and commenting on each other's journals. Additionally, such features should be carefully designed to be more peer-based and conducted within closed groups. Otherwise, as our formative study reflects, direct exposure to the larger community may in turn hinder older adults' participation.

Building on participants' feedback, which highlighted a preference for being "invited to participate," and to better guide older adults during their on-boarding process using a technology support

tool, the Chatbot View could be expanded into a global guidance assistant. The assistant could feature a lively avatar with emotional expressions and voice interactions, which are elements shown to improve older adults' acceptance and engagement in previous studies [20, 80]. Moreover, the assistant could be designed with appropriate proactivity [72] to prompt older adults to write after they finish reading some other people's works in the Square View, and facilitate the communication of editing suggestions in the Edit View, enhancing overall user engagement.

9.5 Limitation and Future Work

First, our goal was to design and develop a tool that provides general support in digital journaling for older adults, where we did not focus on catering to some individual factors, such as differences in existing journaling habits and abilities among older adults. As suggested in our results, these differences may result in different comfort level and perception of LLM assistance between different user groups. Additionally, most participants in our user studies were female, and thus it is unknown whether gender may affect the findings or not. In future work, it would be beneficial to investigate the possible diverse needs resulting from these factors and design more tailored support for specific user groups accordingly.

Second, our work utilized the GPT-4 APIs for implementing the chatbot dialogue and sample text generation features. Although most older adults found that the generated content accurately expressed their thoughts and emotions, there were also instances of inaccuracies like downplaying the negative emotion and sublimated it into an unpleasant but also memorable experience (P9). In future work, more complex frameworks can be proposed to build such LLM-powered systems for older adults' digital journaling. Specific modules can be designed to handle the generation of relatively negative outputs and preserve user's intent and emotions.

Third, although our user study included a field deployment, we did not conduct a more extensive long-term evaluation of JournalAIde. In our deployment study, participants' impression of using JournalAIde was still more like experiencing new technology-assisted writing in a familiar environment. Future research could focus on conducting longitudinal studies to evaluate the sustained support and effectiveness of JournalAIde, exploring changes in user behavior, adaptation patterns, as well as length and content enrichment in writing outputs.

10 Conclusion

In this work, we introduced JournalAIde, a tool powered by large language model (LLM) designed to support digital journaling for older adults. Through a formative study, we identified challenges older adults encounter during digital journaling in terms of confidence, writing ability, and sustained attention, as well as their preferences and needs for technical support. Based on those findings, we derived four design goals to guide the development of JournalAIde. Through a between-subjects user study and a 10-day field deployment study, we demonstrated the effectiveness of JournalAIde on supporting older adults' digital journaling and investigated their user experiences, expectations, and perception of LLM writing assistance. Finally, we discussed the insights and design considerations for future systems aimed to support digital journaling for older adults like JournalAIde.

Acknowledgments

We express gratitude to the reviewers for their insightful suggestions. This work is partially supported by the Guangzhou-HKUST(GZ) Joint Funding Project (No. 2024A03J0617), Guangzhou Higher Education Teaching Quality and Teaching Reform Project (No. 2024YBJG070), Education Bureau of Guangzhou Municipality Funding Project (No. 2024312152), Guangdong Provincial Key Lab of Integrated Communication, Sensing and Computation for Ubiquitous Internet of Things (No. 2023B1212010007), the Project of DEGP (No.2023KCXTD042), and the Guangzhou Science and Technology Program City-University Joint Funding Project (No. 2023A03J0001).

References

[1] Yasser Al-Shboul and Ibrahim Fathi Huwari. 2015. The Causes of Writing Apprehension through Students' Perspective. *Journal of Language Teaching & Research* 6, 3 (2015).
[2] Christina Baldwin. 1977. *One to one: Self-understanding through journal writing.* Rowman & Littlefield.
[3] Christina Baldwin. 1990. *Life's Companion: Journal Writing as a Spiritual Practice.* Bantam.
[4] Albert Bandura and Sebastian Wessels. 1994. *Self-efficacy.* Vol. 4. na.
[5] Ronald W Berkowsky, Joseph Sharit, and Sara J Czaja. 2017. Factors predicting decisions about technology adoption among older adults. *Innovation in aging* 1, 3 (2017), igy002.
[6] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.
[7] Robin Brewer and Anne Marie Piper. 2016. " Tell It Like It Really Is" A Case of Online Content Creation and Sharing Among Older Adult Bloggers. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems.* 5529–5542.
[8] Robin N Brewer and Anne Marie Piper. 2017. xPress: Rethinking design for aging and accessibility through an IVR blogging system. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017), 1–17.
[9] John Brooke. 1996. SUS—a quick and dirty usability scale. 1996. *Usability Eval Ind* 189, 194 (1996), 4–7.
[10] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. *Advances in neural information processing systems* 33 (2020), 1877–1901.
[11] Oliver Burmeister. 2012. What seniors value about online community. *Journal of community informatics* 8, 1 (2012), 1–12.
[12] Jill Burstein, Norbert Elliot, Beata Beigman Klebanov, Nitin Madnani, Diane Napolitano, Maxwell Schwartz, Patrick Houghton, and Hillary Molloy. 2018. Writing mentor: Writing progress using self-regulated writing support. *Journal of Writing Analytics* 2 (2018), 285–313.
[13] Janis M Campbell. 1992. Treating depression in well older adults: Use of diaries in cognitive therapy. *Issues in Mental Health Nursing* 13, 1 (1992), 19–29.
[14] Jessy Ceha, Ken Jen Lee, Elizabeth Nilsen, Joslin Goh, and Edith Law. 2021. Can a humorous conversational agent enhance learning experience and outcomes?. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems.* 1–14.
[15] Montserrat Celdrán, Rodrigo Serrat, and Feliciano Villar. 2019. Older adults as internet content producers: Motivations for blogging in later life. *Perspectives on human-computer interaction research with older people* (2019), 169–182.
[16] Montserrat Celdrán, Rodrigo Serrat, Feliciano Villar, and Roger Montserrat. 2022. Exploring the benefits of proactive participation among adults and older people by writing blogs. *Journal of Gerontological Social Work* 65, 3 (2022), 320–336.
[17] Elizabeth Clark, Anne Spencer Ross, Chenhao Tan, Yangfeng Ji, and Noah A Smith. 2018. Creative writing with a machine in the loop: Case studies on slogans and stories. In *23rd International Conference on Intelligent User Interfaces.* 329–340.
[18] Jacob Cohen. 2013. *Statistical power analysis for the behavioral sciences.* Academic press.
[19] Sara J Czaja, Neil Charness, Arthur D Fisk, Christopher Hertzog, Sankaran N Nair, Wendy A Rogers, and Joseph Sharit. 2006. Factors predicting the use of technology: findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). *Psychology and aging* 21, 2 (2006), 333.
[20] Noemi da Paixão Pinto, Juliana Baptista dos Santos França, Henrique Prado de Sá Sousa, Adriana Santarosa Vivacqua, and Ana Cristina Bicharra Garcia. 2021.

Conversational agents for elderly interaction. In *2021 IEEE 24th international conference on computer supported cooperative work in design (CSCWD)*. IEEE, 1–6.
[21] John A Daly. 1978. Writing apprehension and writing competency. *The journal of educational research* 72, 1 (1978), 10–14.
[22] Edward L Deci and Richard M Ryan. 2008. Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian psychology/Psychologie canadienne* 49, 3 (2008), 182.
[23] Kevin Doherty and Gavin Doherty. 2018. Engagement in HCI: conception, theory and measurement. *ACM Computing Surveys (CSUR)* 51, 5 (2018), 1–39.
[24] Lester Faigley and Stephen Witte. 1981. Analyzing revision. *College Composition & Communication* 32, 4 (1981), 400–414.
[25] Mingming Fan and Khai N Truong. 2018. Guidelines for creating senior-friendly product instructions. *ACM Transactions on Accessible Computing (TACCESS)* 11, 2 (2018), 1–35.
[26] Marcelo Fernández, Iyubanit Rodríguez, Pedro O Rossel, Carolina Fuentes, and Valeria Herskovic. 2017. InMyDay: a digital diary to promote self-care among elders. In *Ubiquitous Computing and Ambient Intelligence: 11th International Conference, UCAmI 2017, Philadelphia, PA, USA, November 7–10, 2017, Proceedings*. Springer, 486–497.
[27] Susan M Ferreira, Sergio Sayago, and Josep Blat. 2017. Older people's production and appropriation of digital videos: an ethnographic study. *Behaviour & Information Technology* 36, 6 (2017), 557–574.
[28] Meredith Flood and Kenneth D Phillips. 2007. Creativity in older adults: A plethora of possibilities. *Issues in mental health nursing* 28, 4 (2007), 389–411.
[29] Linda Flower and John R Hayes. 1981. A cognitive process theory of writing. *College composition and communication* 32, 4 (1981), 365–387.
[30] Catherine O Fritz, Peter E Morris, and Jennifer J Richler. 2012. Effect size estimates: current use, calculations, and interpretation. *Journal of experimental psychology: General* 141, 1 (2012), 2.
[31] Nancy C Goodwin. 1987. Functionality and usability. *Commun. ACM* 30, 3 (1987), 229–233.
[32] Keith N Hampton, Lauren Sessions Goulet, Lee Rainie, and Kristen Purcell. 2011. *Social networking sites and our lives*. Vol. 1. Pew Internet & American Life Project Washington, DC.
[33] Sophie M Hardy, Katrien Segaert, and Linda Wheeldon. 2020. Healthy aging and sentence production: Disrupted lexical access in the context of intact syntactic planning. *Frontiers in Psychology* 11 (2020), 257.
[34] Dave Harley and Geraldine Fitzpatrick. 2009. YouTube and intergenerational communication: the case of Geriatric1927. *Universal access in the information society* 8 (2009), 5–20.
[35] Lynn Hasher and Rose T Zacks. 1988. Working memory, comprehension, and aging: A review and a new view. *Psychology of learning and motivation* 22 (1988), 193–225.
[36] Simone Hausknecht, Michelle Vanchu-Orosco, and David Kaufman. 2019. Digitising the wisdom of our elders: Connectedness through digital storytelling. *Ageing & Society* 39, 12 (2019), 2714–2734.
[37] Roger Hiemstra et al. 2001. Uses and benefits of journal writing. *New directions for adult and continuing education* 2001, 90 (2001), 19.
[38] Axel Hoesl and Andreas Butz. 2017. Sense of Authorship and Agency in Computational Creativity Support.. In *MICI@ CHI*.
[39] Christine Hogan. 1995. Creative and reflective journal processes. *The Learning Organization* 2, 2 (1995), 4–17.
[40] Maureen Hoskyn and H Lee Swanson. 2003. The relationship between working memory and writing in younger and older adults. *Reading and Writing* 16 (2003), 759–784.
[41] Ziwei Ji, YU Tiezheng, Yan Xu, Nayeon Lee, Etsuko Ishii, and Pascale Fung. 2023. Towards mitigating LLM hallucination via self reflection. In *The 2023 Conference on Empirical Methods in Natural Language Processing*.
[42] Xiaofu Jin, Emily Kuang, and Mingming Fan. 2021. " Too old to bank digitally?": A Survey of Banking Practices and Challenges Among Older Adults in China. In *Designing Interactive Systems Conference 2021*. 802–814.
[43] Daniel Jurafsky and James H. Martin. 2024. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models* (3rd ed.). https://web.stanford.edu/~jurafsky/slp3/ Online manuscript released August 20, 2024.
[44] Youwen Kang, Zhida Sun, Sitong Wang, Zeyu Huang, Ziming Wu, and Xiaojuan Ma. 2021. MetaMap: Supporting visual metaphor ideation through multi-dimensional example-based exploration. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–15.
[45] Julia Karbach and Paul Verhaeghen. 2014. Making working memory work: A meta-analysis of executive-control and working memory training in older adults. *Psychological science* 25, 11 (2014), 2027–2037.
[46] Francis E Kazemek. 1999. " A gathering of individuals": A longitudinal study of a writing workshop for older adults. *Adult Basic Education* 9, 1 (1999), 3.
[47] Shoshi Keisari, Silvia Piol, Talia Elkarif, Giada Mola, and Ines Testoni. 2021. Crafting life stories in photocollage: An online creative art-based intervention for older adults. *Behavioral Sciences* 12, 1 (2021), 1.
[48] Seoyoung Kim, Donghoon Shin, Jeongyeon Kim, Soonwoo Kwon, and Juho Kim. 2023. How Older Adults Use Online Videos for Learning. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–16.
[49] Taewan Kim, Seolyeong Bae, Hyun Ah Kim, Su-Woo Lee, Hwajung Hong, Chanmo Yang, and Young-Ho Kim. 2024. MindfulDiary: Harnessing Large Language Model to Support Psychiatric Patients' Journaling. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 701, 20 pages. doi:10.1145/3613904.3642937
[50] Taewan Kim, Donghoon Shin, Young-Ho Kim, and Hwajung Hong. 2024. DiaryMate: Understanding User Perceptions and Experience in Human-AI Collaboration for Personal Journaling. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 1046, 15 pages. doi:10.1145/3613904.3642693
[51] Heidi E Koschwanez, Ngaire Kerse, Margot Darragh, Paul Jarrett, Roger J Booth, and Elizabeth Broadbent. 2013. Expressive writing and wound healing in older adults: a randomized controlled trial. *Psychosomatic Medicine* 75, 6 (2013), 581–590.
[52] Amanda Lazar, Mark Diaz, Robin Brewer, Chelsea Kim, and Anne Marie Piper. 2017. Going gray, failure to hire, and the ick factor: Analyzing how older bloggers talk about ageism. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. 655–668.
[53] Omer Levy and Yoav Goldberg. 2014. Linguistic regularities in sparse and explicit word representations. In *Proceedings of the eighteenth conference on computational natural language learning*. 171–180.
[54] Hsiu-Li Liao, Su-Houn Liu, and Shih-Ming Pi. 2011. Modeling motivations for blogging: An expectancy theory analysis. *Social Behavior and Personality: an international journal* 39, 2 (2011), 251–264.
[55] Chengzhong Liu, Zeyu Huang, Dingdong Liu, Shixu Zhou, Zhenhui Peng, and Xiaojuan Ma. 2022. PlanHelper: Supporting Activity Plan Construction with Answer Posts in Community-based QA Platforms. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–26.
[56] Chengzhong Liu, Shixu Zhou, Dingdong Liu, Junze Li, Zeyu Huang, and Xiaojuan Ma. 2023. CoArgue: Fostering Lurkers' Contribution to Collective Arguments in Community-based QA Platforms. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–17.
[57] Brianna O Malyn, Zoe Thomas, and Christine E Ramsey-Wade. 2020. Reading and writing for well-being: A qualitative exploration of the therapeutic experience of older adult participants in a bibliotherapy and creative writing group. *Counselling and Psychotherapy Research* 20, 4 (2020), 715–724.
[58] Potsawee Manakul, Adian Liusie, and Mark JF Gales. 2023. Selfcheckgpt: Zero-resource black-box hallucination detection for generative large language models. *arXiv preprint arXiv:2303.08896* (2023).
[59] Henry B Mann and Donald R Whitney. 1947. On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics* (1947), 50–60.
[60] Courtney McAlister and Maureen Schmitter-Edgecombe. 2016. Cross-sectional and longitudinal analyses of everyday memory lapses in older adults. *Aging, Neuropsychology, and Cognition* 23, 5 (2016), 591–608.
[61] E Michael Brady and Harry Z Sky. 2003. Journal writing among older learners. *Educational Gerontology* 29, 2 (2003), 151–163.
[62] David R Millen. 2000. Rapid ethnography: time deepening strategies for HCI field research. In *Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques*. 280–286.
[63] Tracy L Mitzner, Jyoti Savla, Walter R Boot, Joseph Sharit, Neil Charness, Sara J Czaja, and Wendy A Rogers. 2019. Technology adoption by older adults: Findings from the PRISM trial. *The Gerontologist* 59, 1 (2019), 34–44.
[64] Bonnie A Nardi, Diane J Schiano, Michelle Gumbrecht, and Luke Swartz. 2004. Why we blog. *Commun. ACM* 47, 12 (2004), 41–46.
[65] Novia Nurain and Chia-Fang Chung. 2023. "I left my legacy, told my story": Understanding Older Adults' Tracking Practices to Promote Active Aging. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference* (Pittsburgh, PA, USA) *(DIS '23)*. Association for Computing Machinery, New York, NY, USA, 459–475. doi:10.1145/3563657.3596083
[66] Heather L O'Brien and Elaine G Toms. 2008. What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American society for Information Science and Technology* 59, 6 (2008), 938–955.
[67] William Odom, Richard Banks, David Kirk, Richard Harper, Siân Lindley, and Abigail Sellen. 2012. Technology heirlooms? considerations for passing down and inheriting digital materials. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 337–346. doi:10.1145/2207676.2207723
[68] Heather O'Brien. 2016. Theoretical perspectives on user engagement. *Why engagement matters: Cross-disciplinary perspectives of user engagement in digital media* (2016), 1–26.
[69] Annlia Paganini-Hill and Linda J. Clark. 2006. Preliminary Assessment of Cognitive Function in Older Adults by Clock Drawing, Box Copying

and Narrative Writing. *Dementia and Geriatric Cognitive Disorders* 23, 2 (11 2006), 74–81. doi:10.1159/000097097 arXiv:https://karger.com/dem/article-pdf/23/2/74/2565583/000097097.pdf
[70] Frank Pajares and Margaret J Johnson. 1994. Confidence and competence in writing: The role of self-efficacy, outcome expectancy, and apprehension. *Research in the Teaching of English* (1994), 313–331.
[71] Zhenhui Peng, Qingyu Guo, Ka Wing Tsang, and Xiaojuan Ma. 2020. Exploring the effects of technological writing assistance for support providers in online mental health community. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–15.
[72] Zhenhui Peng, Yunhwan Kwon, Jiaan Lu, Ziming Wu, and Xiaojuan Ma. 2019. Design and evaluation of service robot's proactivity in decision-making support process. In *proceedings of the 2019 CHI conference on human factors in computing systems*. 1–13.
[73] Rebecca Pera, Sarah Quinton, and Gabriele Baima. 2020. I am who I am: Sharing photos on social media by older consumers and its influence on subjective well-being. *Psychology & Marketing* 37, 6 (2020), 782–795.
[74] Ira Progoff. 1977. *At a journal workshop: The basic text and guide for using the Intensive Journal.* Dialogue House Library.
[75] EHUD REITER and ROBERT DALE. 1997. Building applied natural language generation systems. *Natural Language Engineering* 3, 1 (1997), 57–87. doi:10.1017/S1351324997001502
[76] Arlind Reuter, Tom Bartindale, Kellie Morrissey, Thomas Scharf, and Jennifer Liddle. 2019. Older voices: Supporting community radio production for civic participation in later life. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–13.
[77] Arlind Reuter and Jennifer Liddle. 2020. The Later Life Audio and Radio Co-operative: considering radio as a technology to promote citizen dialogue in later life. In *Proceedings of the 9th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion*. 153–157.
[78] Arlind Reuter, Thomas Scharf, and Jan Smeddinck. 2021. Content Creation in Later Life: Reconsidering Older Adults' Digital Participation and Inclusion. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW3, Article 257 (jan 2021), 23 pages. doi:10.1145/3434166
[79] Valeria Righi, Sergio Sayago, and Josep Blat. 2017. When we talk about older people in HCI, who are we talking about? Towards a 'turn to community'in the design of technologies for a growing ageing population. *International Journal of Human-Computer Studies* 108 (2017), 15–31.
[80] Antonia Rodríguez-Martínez, Teresa Amezcua-Aguilar, Javier Cortés-Moreno, and Juan José Jiménez-Delgado. 2023. Qualitative Analysis of Conversational Chatbots to Alleviate Loneliness in Older Adults as a Strategy for Emotional Health. In *Healthcare*, Vol. 12. MDPI, 62.
[81] Sergio Sayago, Paula Forbes, and Josep Blat. 2012. Older people's social sharing practices in YouTube through an ethnographical lens. In *The 26th BCS Conference on Human Computer Interaction 26*. 185–194.
[82] Anthony Seow. 2002. The writing process and process writing. *Methodology in language teaching: An anthology of current practice* 315 (2002), 320.
[83] Craig E Shepherd and Steven Aagard. 2011. Journal writing with web 2.0 tools: a vision for older adults. *Educational Gerontology* 37, 7 (2011), 606–620.
[84] Xinru Tang, Yuling Sun, Bowen Zhang, Zimi Liu, RAY Lc, Zhicong Lu, and Xin Tong. 2022. " I Never Imagined Grandma Could Do So Well with Technology" Evolving Roles of Younger Family Members in Older Adults' Technology Learning and Use. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–29.
[85] Maciej Tomczak and Ewa Tomczak. 2014. The need to report effect size estimates revisited. An overview of some recommended measures of effect size. (2014).
[86] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *Advances in neural information processing systems* 30 (2017).
[87] Kerryellen G Vroman, Sajay Arthanat, and Catherine Lysack. 2015. "Who over 65 is online?" Older adults' dispositions toward information communication technology. *Computers in Human Behavior* 43 (2015), 156–166.
[88] Jenny Waycott, Frank Vetere, Sonja Pedell, Lars Kulik, Elizabeth Ozanne, Alan Gruner, and John Downs. 2013. Older adults as digital content producers. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 39–48.
[89] Stephan Weibelzahl, Alexandros Paramythis, and Judith Masthoff. 2020. Evaluation of adaptive systems. In *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and personalization*. 394–395.
[90] Rainer Winkler, Sebastian Hobert, Antti Salovaara, Matthias Söllner, and Jan Marco Leimeister. 2020. Sara, the lecturer: Improving learning in online education with a scaffolding-based conversational agent. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. 1–14.
[91] Sam Wiseman, Stuart M Shieber, and Alexander M Rush. 2017. Challenges in data-to-document generation. *arXiv preprint arXiv:1707.08052* (2017).
[92] Meng Xia, Qian Zhu, Xingbo Wang, Fei Nie, Huamin Qu, and Xiaojuan Ma. 2022. Persua: A visual interactive system to enhance the persuasiveness of arguments in online discussion. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–30.
[93] Ann Yuan, Andy Coenen, Emily Reif, and Daphne Ippolito. 2022. Wordcraft: story writing with large language models. In *27th International Conference on Intelligent User Interfaces*. 841–852.
[94] Salifu Yusif, Jeffrey Soar, and Abdul Hafeez-Baig. 2016. Older people, assistive technologies, and the barriers to adoption: A systematic review. *International journal of medical informatics* 94 (2016), 112–116.
[95] Yue Zhang, Yafu Li, Leyang Cui, Deng Cai, Lemao Liu, Tingchen Fu, Xinting Huang, Enbo Zhao, Yu Zhang, Yulong Chen, et al. 2023. Siren's song in the AI ocean: a survey on hallucination in large language models. *arXiv preprint arXiv:2309.01219* (2023).
[96] Zhe Victor Zhang and Ken Hyland. 2018. Student engagement with teacher and automated feedback on L2 writing. *Assessing Writing* 36 (2018), 90–102.
[97] Wei Zhao, Ryan M Kelly, Melissa J Rogerson, and Jenny Waycott. 2023. Older Adults Using Technology for Meaningful Activities During COVID-19: An Analysis Through the Lens of Self-Determination Theory. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–17.

A Prompts for Calling LLM

A.1 Chatbot Conversation

You are an assistant who helps the older adults write digital journals. You should ask questions based on the information provided by the older adults to help them recall and talk about their experience and ideas. Please note: First, please only ask one question at a time. Second, the questions need to be short and uncomplicated. Third, use a gentle tone to encourage the older adults to recall and think.

A.2 Text Regeneration

You are an assistant who helps the older adults write digital journals. Next you will receive a piece of text and the older adult's modification ideas. Please make modifications based on the older adult's ideas.

A.3 Consistency Scoring

You are an assistant who helps evaluate the consistency between the generated text and user's description. I will give you the user's description and a generated sentence. Please use a decimal number between 0 and 1 to answer the consistency between the generated sentence and the user's description.
User's description: <User Input Information>
Generated sentence: <Generated Sentence>

B Original System Interface

The original system interfaces are shown in Figure 9 & 10. The translation strategy used to produce the figures in the Section 4 is Google Translate.

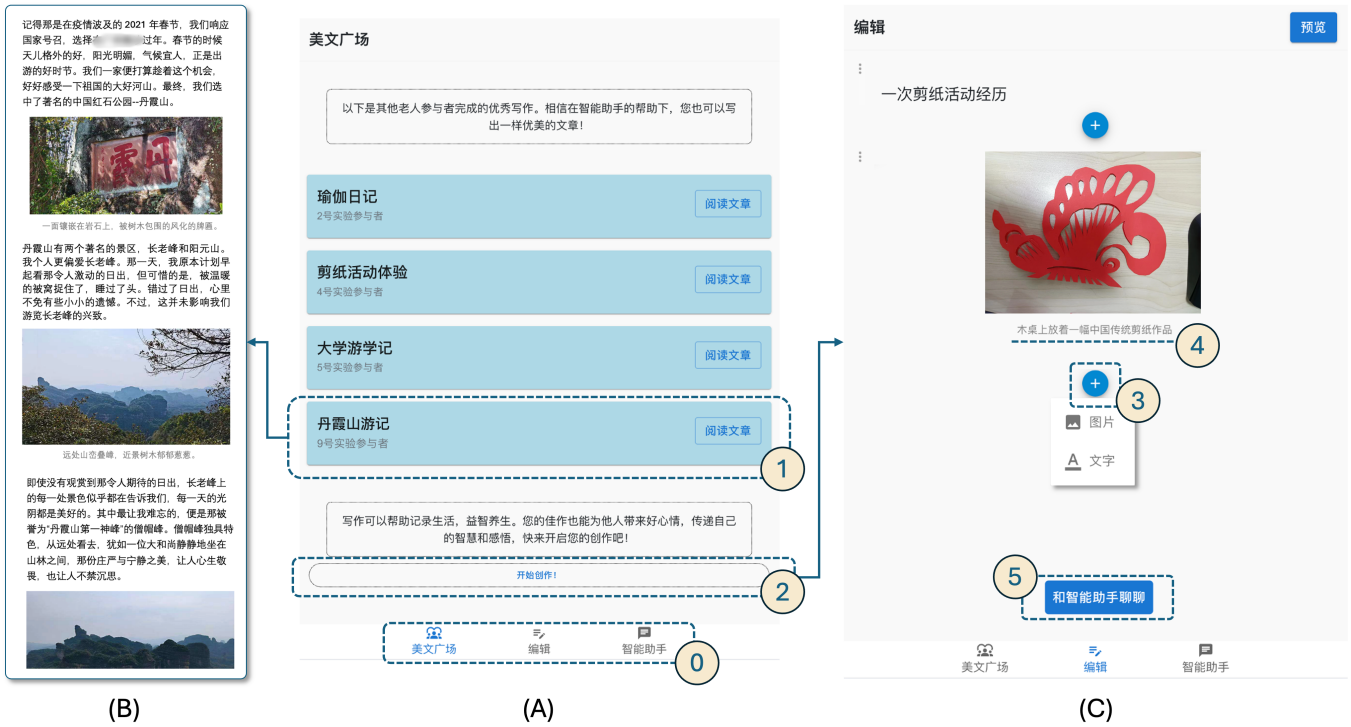


Figure 9: Original interface of Figure 2

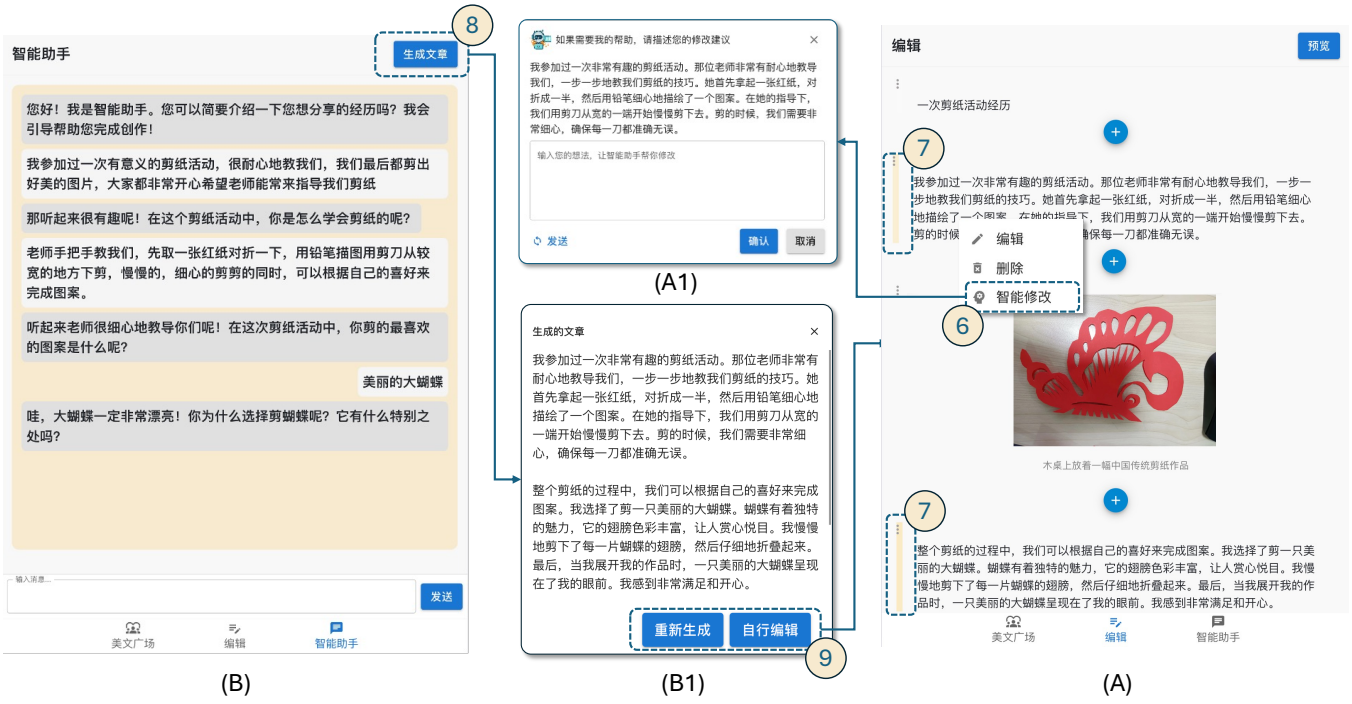


Figure 10: Original interface of Figure 3