# ExploreSelf: Fostering User-driven Exploration and Reflection on Personal Challenges with Adaptive Guidance by Large Language Models

Inhwa Song*
School of Computing
KAIST
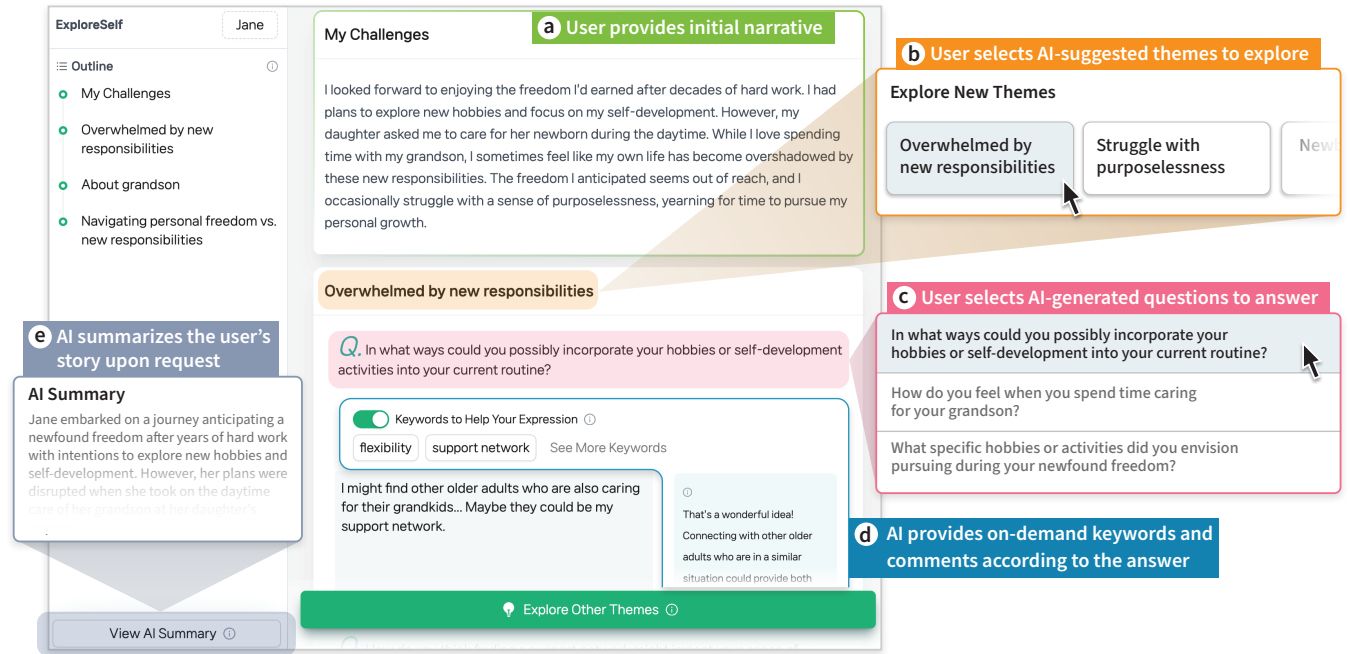Daejeon, Republic of Korea
inhwa.song@kaist.ac.kr

SoHyun Park
NAVER Cloud
Seongnam, Republic of Korea
sohyun@snu.ac.kr

Sachin R Pendse
Northwestern University
Evanston, Illinois, USA
sachin.r.pendse@northwestern.edu

Jessica Lee Schleider
Medical Social Sciences, Pediatrics,
and Psychology
Northwestern University
Evanston, Illinois, USA
jessica.schleider@northwestern.edu

Munmun De Choudhury
School of Interactive Computing
Georgia Institute of Technology
Atlanta, Georgia, USA
munmund@gatech.edu

Young-Ho Kim
NAVER AI Lab
Seongnam, Gyeonggi, Republic of
Korea
yghokim@younghokim.net

Figure 1: ExploreSelf supports user-driven exploration and reflection on their personal challenges. Based on the initial narrative ⓐ written by the user, they can explore diverse perspectives by following up the AI-suggested or custom themes ⓑ. Regarding each theme, the user can reflect on detailed aspects through questions generated by an AI ⓒ. As adaptive guidance, the AI provides keywords and comments on the fly that the user can refer to while writing an answer ⓓ. At any time, the user can request an AI-generated summary of the current exploration ⓔ.

## Abstract

Expressing stressful experiences in words is proven to improve mental and physical health, but individuals often disengage with writing interventions as they struggle to organize their thoughts

and emotions. Reflective prompts have been used to provide direction, and large language models (LLMs) have demonstrated the potential to provide tailored guidance. However, current systems often limit users' flexibility to direct their reflections. We thus present ExploreSelf, an LLM-driven application designed to empower users to control their reflective journey, providing adaptive support through dynamically generated questions. Through an exploratory study with 19 participants, we examine how participants explore and reflect on personal challenges using ExploreSelf. Our findings demonstrate that participants valued the flexible navigation of adaptive guidance to control their reflective journey, leading to deeper engagement and insight. Building on our findings, we discuss the implications of designing LLM-driven tools that facilitate user-driven and effective reflection of personal challenges.

## CCS Concepts

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

## Keywords

Reflective Writing, Technology for Well-being, User-driven Exploration, User Agency, Large Language Models

## 1 Introduction

Articulating stressful or emotional experiences in words has been found to improve both physical and mental health by enabling individuals to process and express their thoughts and emotions [6, 88, 89]. It has been a widely circulated and popular means for individuals to self-reflect and make sense of themselves [27, 39], by simply sitting down to write the innermost thoughts and feelings without confines of grammar, punctuation, or any other formalities of written composition. However, writing about negative experiences and thoughts can be challenging as it often requires individuals to confront distressing thoughts and emotions, making it tempting for them to disengage from the process [89, 110]. Hence, there have been a number of adjustments to the activity to help with this process. For example, tools like self-help workbooks and prompted journaling have been developed to offer structured prompts that guide users in exploring their thoughts and emotions [74, 95]. The HCI community has also introduced interactive methods, such as using photos as journaling prompts [43] or employing a conversational assistant with sequences of probing questions [83]. These approaches mostly involved pre-defined prompts for writers to use, which has proven effective to spark thinking and dive into the process [21, 45]. It has also been noted that these prompts are inflexible and detached from the user context [9].

The generative capabilities of recent large language models (LLMs) have accelerated the advancement of intelligent writing support tools [55, 56], which commonly perform natural language generation for directly modifying the content (*e.g.*, [48, 56]) or providing advice for improving the content (*e.g.*, [28]). However, research on LLM-driven writing tools largely focuses on creative writing, and only a handful of works have recently begun to study intelligent support for writing about the self (*e.g.*, DiaryMate [48]). Further, a growing body of research has investigated the LLMs' potential for mental health support in the form of conventional chat interfaces (*e.g.*, [47, 52, 105, 113]). Although LLM-driven, open-ended conversations are intuitively analogous to therapists' counseling sessions, an agent who holds the therapist's persona tends to lead the conversation and sometimes can result in unintentional reinforcement of negative narratives [40, 47, 113]. Meanwhile, research suggests that when individuals are able to make choices and exert control over their experiences, they engage more meaningfully in the process and develop a stronger sense of ownership and responsibility for their growth [23].

In this work, we investigate ways to empower individuals to take greater control of the reflective process of their personal challenges while simultaneously leveraging the personalized and adaptive guidance offered by LLMs. Unlike traditional mental health services that focus on professional diagnosis and intervention, our goal was to design a system that enables individuals to explore their challenges at their own pace, without clinical oversight. To this aim, we designed and developed ExploreSelf (Figure 1), an LLM-infused interactive system for exploring and reflecting on personal challenges. ExploreSelf allows users to freely expand related themes from the initial narrative and drill down into probing questions to reflect on aspects of themes in depth. To facilitate exploration, the system provides on-demand guidance, such as prompting keywords and comments tailored to their writing. The design of ExploreSelf was informed by formative interviews with nine mental health professionals, where their challenges and approaches to managing the counseling sessions inspired our interface design and the generative pipelines.

To understand how people interact and engage with ExploreSelf in exploring their personal challenges, we conducted an online lab study with 19 individuals, where they explored their personal challenges using ExploreSelf for 30 to 60 minutes. Through user-directed exploration with diverse forms of adaptive guidance by LLMs, ExploreSelf enabled individuals to engage in meaningful self-reflection while maintaining control over the direction of their reflection. Participants actively leveraged the navigation components (*i.e.*, themes, questions, and summary) and the adaptive guidance (*i.e.*, keywords, and comments) in individualized patterns. Participants' perceived agency was significantly increased after the session. Participants reported that the system's adaptive features, such as the ability to select themes, choose from multiple Socratic questions, and toggle keywords and comments, contributed to their sense of control and autonomy. Participants decided which aspects to focus on or revisit later in a way that balances their cognitive and emotional effort with the expected significance of the engagement regarding their personal concerns. Drawing on the findings, we discuss how the components of ExploreSelf collectively fostered a dynamic and personalized reflective process, helping people feel more empowered in navigating their personal challenges.

The key contributions of this work are threefold:

(1) The design and implementation of EXPLORESELF, a novel LLM-driven web application that supports guided writing about breakdown topics regarding a personal issue, encouraging sensemaking and refreshment of thoughts on it. EXPLORESELF's design was informed by formative interviews ($N = 9$) with clinical psychologists and professional licensed counselors. The source code of EXPLORESELF is publicly available at https://naver-ai.github.io/exploreself.

(2) An empirical study using EXPLORESELF, with 19 individuals with varied severity and types of personal issues. From the quantitative and qualitative analysis, we provide an understanding of how people elaborate thoughts and make sense of their own issues through interactive reflection aided by adaptive guidance. We uncover how participants balanced exploration and reflection, engaging with AI-driven themes, questions, keywords, and comments, and offer insights into their perspectives on these forms of guidance.

(3) Implications for designing LLM-guided tools that facilitate user-driven and effective reflection of personal challenges, such as balancing user autonomy with guided support, fostering long-term practice, and implementing the nuanced role and stance of AI to accommodate a diverse range of individuals.

## 2 Related Work

In this section, we cover the related work in the areas of (1) writing for well-being and (2) technology-mediated writing for well-being.

### 2.1 Writing for Well-being

A substantial body of research has demonstrated that the act of writing as an instrument for exploring personal thoughts and emotions can promote self-healing and personal growth, significantly enhancing both physical and emotional well-being [6, 88, 89]. It has been widely studied across diverse contexts, including expressive writing from social psychology [89, 98], integration into specific psychotherapies as a treatment (*e.g.*, guided autobiography, using diary in CBT), positive interventions to promote psychological well-being (*e.g.*, writing about gratitude and forgiveness), and through personal practices [98]. One of the most widely researched methods is expressive writing, which involves writing about one's deepest thoughts and feelings regarding stressful experiences [89]. Across these various writing for well-being practices, the common focus lies in translating personal experiences into written language, allowing individuals to process and organize their thoughts and feelings in a meaningful way [88–90, 98].

Through this process, individuals often derive meaning from stressful events and foster a sense of hope as they work through their emotions and gain new perspectives [34]. This activity has been shown to yield significant benefits for individuals across various demographics, including different age groups [51, 117], gender [66], cross-cultures [25, 49], and health conditions [63, 97]. Studies have also found that writing can be beneficial for well-being when people engage in writing outside of controlled laboratory settings, such as at home or online [60]. While several theories attempt to explain the connection between expressive writing and health, no single mechanism has emerged as dominant, likely due to the

varied experiences and expectations of participants [54]. However, studies from psychology have suggested that writing allows individuals to release suppressed emotions [59], gain deeper insights into their experiences [88], and develop a sense of control over their thoughts and emotions through self-reflection and emotion regulation [59]. In this way, writing can be used intentionally and purposefully to foster self-discovery, emotional growth, and personal insight, enabling individuals to gain a deeper understanding of themselves and their experiences [1, 94, 116].

Research from psychology has indicated that writing about stressful events demands a significant level of perceived control, as it helps individuals organize, connect, and contextualize their thoughts and emotions [3, 31]. This sense of control is not only essential for initiating or sustaining a writing practice but is also closely tied to physical and psychological well-being, making it a core component of the writing process itself [3]. To support individuals in maintaining a sense of control, structured guidance has been studied as an effective tool to reduce cognitive load and alleviate the overwhelm that often accompanies reflection. Prompts and reflection frameworks are amongst the popularly studied tools for offering that sense of control, offering a sense of direction or goal to writers. In this way, individuals can navigate their thoughts and emotions in a more organized way, providing manageable steps to follow rather than leaving them to explore themselves entirely on their own. This approach supports sustained engagement by making the process more approachable and less overwhelming [8, 19]. Previous research has explored various methods for designing such structure, including gratitude-focused writing [29], writing in a narrative structure [20], using sentence stems (*e.g.*, "*The thing I am most worried about is...*"), journaling with photographs [76], and mind maps [20, 116]. While these approaches offer helpful starting points, they are often provided in a fixed format, usually requiring the individual to stay on a structured course throughout, therefore they cannot respond to the writer's thoughts and feelings that evolve in writing. This limits their ability to provide adaptive feedback or tailored suggestions to the individual's unique emotional states and goals in writing. Recognizing this gap, recent works have explored more interactive methods that provide personalized prompts, real-time feedback, and adaptive support [41, 72, 118].

### 2.2 Technology-Mediated Writing for Well-being

With the soaring number of digital writing tools, individuals increasingly engage in writing for well-being using these tools, such as on online health platforms and digital journaling tools, to motivate themselves to engage in writing for well-being. [64, 111]. These spontaneous, technology-mediated practices provide a space for self-expression and reflection, yet users often face challenges in maintaining consistent engagement and fostering deeper introspection. In response, the HCI community has explored methods to support individuals in their reflective writing journeys, focusing on strategies to enhance engagement by introducing the social presence of chatbots [83] and providing contextual data to aid reflection [37]. These approaches have aimed to create structured yet flexible systems that guide users in navigating their thoughts and emotions more effectively.

The generative capabilities of LLMs (*e.g.*, GPT [12], Gemini [115], Claude [4], HyperCLOVA [46]) have motivated the development of flexible and adaptive support for writing for well-being (*e.g.*,[47, 48, 77, 107, 121]), offering personalized interactions that acknowledge and are responsive to what users write. For example, Kim *et al.* explored the potential of interactive dialogue with LLMs as an alternative format for reflective journaling for psychiatric patients, where LLM chatbots provided adaptive reflective questions [47]. Their findings demonstrated that these adaptive prompts benefited individuals by helping them explore and reflect on their past experiences and emotions. However, since interactions with LLMs are naturally and intuitively conversational in nature, many LLM-driven approaches let users "chat" with an AI agent rather than write [40, 47, 81, 113]. Conventional turn-taking structure in chat inherently confines users to respond to the agent's last message, and such linearly progressing dialogues limit the flexibility to pivot subjects or grasp the landscape of the conversation. This makes the user's thought process heavily influenced by the performance and behavioral design of the AI agent, introducing a risk of triggering negative reinforcement [14, 65, 84, 122].

Research suggests that when individuals are given the freedom to make choices and exercise control over their reflective processes, they tend to experience deeper engagement driven by intrinsic motivation [85]. This autonomy fosters a stronger sense of ownership over their personal growth, increasing their responsibility for and investment in the outcomes of their reflection [23, 26, 68, 85, 99]. Recently in HCI, self-guided tools have been studied in the context of therapeutic intervention (*e.g.*, [106, 107]), allowing individuals to engage with LLM-driven systems to reframe negative patterns of thinking. These systems aim to improve the accessibility and effectiveness of interventions by guiding users through well-established tasks, such as identifying cognitive distortions and reframing, with the LLM providing adaptive support throughout the cognitively and emotionally demanding process. However, while such predefined therapeutic modules provide focused support, in this case on correcting cognitive distortions, these systems may still keep the users from exploring the freely flowing thoughts and feelings of users and coming to a resolution or new self-insight on their own. In this work, we are interested in designing an LLM-driven writing for well-being activity that can preserve user agency throughout the reflection process while still providing adaptive support. Striking this balance between guidance and autonomy is crucial, and we turn to formative studies where we explore ways to integrate adaptive LLM support within the context of users' reflective journeys.

## 3 Formative Interviews

To inform the design of ExploreSelf, we conducted semi-structured interviews with mental health professionals, whose main role is to guide individuals in exploring and addressing their personal challenges. By understanding their underlying strategies and expertise, we aimed to gain a nuanced understanding of the challenges of providing effective guidance, which in turn informs the key considerations for designing an interface that incorporates therapeutic principles for its guidance, as well as to obtain insights into the potential and proper role of AI.

**Table 1: Demographic information of the experts who participated in our formative study.**

| Alias | Gender | Job title | Years of experience | Country of experience |
|---|---|---|---|---|
| E1 | Female | Licensed counselor | 16 years | South Korea |
| E2 | Female | Clinical psychologist | 26 years | South Korea |
| E3 | Female | Licensed counselor | 15 years | South Korea |
| E4 | Female | Licensed counselor | 17 years | South Korea |
| E5 | Female | Psychiatrist | 18 years | South Korea |
| E6 | Female | Clinical psychologist | 20 years | United States, United Kingdom |
| E7 | Female | Clinical psychologist | 11 years | United States |
| E8 | Male | Clinical psychologist | 5 years | United States |
| E9 | Female | Clinical psychologist | 8 years | United States |

### 3.1 Procedure and Analysis

We recruited nine experts (E1–9, eight females and one male; see Table 1) through snowball sampling and internal network, who have extensive experience in clinical psychology and professional counseling. We initially planned to recruit experts in South Korea as our system was intended to be evaluated with the Korean population. However, the prolonged medical crisis in South Korea in 2024 made it difficult to recruit clinical psychologists locally. To supplement expert feedback from clinical psychologists, we recruited four clinical psychologists from outside South Korea through our internal network. The final nine participants comprised five clinical psychologists, three licensed therapists, and one psychiatrist. Five were based in South Korea, three in the United States, and one had experience working across the United States and the United Kingdom. Participants had an average of 15.11 years (ranged 5 to 26) of experience in mental health care. Each interview lasted about one hour and was conducted in person or remotely, depending on the participants' availability. We compensated participants with 100,000 KRW or 100 USD, depending on their residence.

The interviews covered participants' strategies for helping clients navigate personal challenges, typical difficulties clients face, and potential roles they envision for LLM-driven systems in the therapeutic process. Since the experts have cared for various client populations in different settings, we asked them to share insights applicable to the general public. To elicit their feedback, we also presented the experts with an example personal challenge narrative (see Appendix A.1) and asked them to describe how they would approach such a case, including how they might guide the client through the process. Three researchers devised the example narrative during a brainstorming session, focusing on creating culturally neutral challenges. Additionally, we presented a paper prototype of an LLM-driven guided writing where the system suggests related topics based on the same example personal challenge narrative. We asked for participants' feedback on its potential utility and areas for improvement.

All interviews were audio-recorded and transcribed for analysis. Employing thematic analysis [11], one researcher coded transcripts and grouped them into broader themes. The research team iterated

several rounds of discussion to refine themes. In the following, we cover the findings from the interviews.

## 3.2 Findings

***Importance of Guided Exploration in Addressing Individual Needs.*** Participants highlighted the essential role of therapists in tailoring guidance to explore the unique needs of clients. E8 noted, "*When clients first enter therapy, they typically have a vague desire to 'feel better' but struggle to articulate specific goals.*" To address such ambiguity, therapists often start conversations to help clients clarify their objectives and understand the underlying issues they need to address. This process not only sets effective goals for the therapy but also empowers clients to gain a clearer understanding of their own difficulties. To guide their clients, all participants commonly used a strategy of asking a series of focused, open-ended questions to encourage self-reflection. E2 shared that she often encourages clients to write about their issues between sessions, but many times, they struggle to decide what topic to write about. E2 remarked, "*I might suggest they write about certain things that arise from our collaborative exploration and give them specific assignments.*"

***Challenges in Effective Questioning during Conversations.*** All participants highlighted the challenges of asking adequate questions that the client would be ready to accept and explore. E6 noted, "*If I ask certain questions too early, the client might push back, thinking I'm making assumptions,*" emphasizing that clients should feel understood and supported, rather than overwhelmed or led in a direction they are not prepared to explore. Similarly, E4 mentioned, "*My biggest challenge is with clients who don't have much to say—they may shut down or resist answering. In these cases, I often have to find alternative questions or pathways to re-engage them.*" Participants pointed out that there is no holy grail when it comes to the sequence of questions, and therefore, the effectiveness of any therapeutic approach depends largely on the therapist's ability to decide when and what the client would find it acceptable.

***Guidance to Enhance Awareness of the Process.*** Participants emphasized the importance of keeping clients informed about the the status of the current therapeutic conversation, in part to stay focused and to prevent them from feeling overwhelmed or lost in their thoughts. They also noted that understanding the flow of the session and the broader discussions context helps clients construct coherent narratives and stay engaged in their own problems. For those struggling with ruminative thinking, where thoughts can circle back repeatedly without recognizing the broader context, these reminders become particularly important: E1 noted, "*They often don't realize they're going in circles until I reflect back on what they have said, which helps them recognize the pattern,*". E4 mentioned that she uses visual aids during sessions by writing down key themes or keywords as the conversation progresses.

## 4 EXPLORESELF

In this section, we first cover the design rationales we derived from prior literature and the formative study. We then describe the system design and user interfaces of EXPLORESELF and implementation details, including the generative pipelines.

## 4.1 Design Rationales

***DR1. Provide diverse pathways to foster user agency and deep reflection.*** Recognizing that individuals may explore their thoughts through various cognitive pathways, it is crucial to foster flexible thinking and support a sense of autonomy in the reflective process [22, 44, 108]. In our formative study, professionals emphasized the importance of allowing clients to control their own reflective journeys. Therapists adjust their questions based on the client's responses and collaboratively identify topics to prioritize for deeper exploration. Therefore, we designed EXPLORESELF to provide a rich resource that enables users to take the lead in shaping and guiding their reflective pathways. First, the EXPLORESELF generates "Themes"—potential areas of reflection—based on the initial narrative provided by the user. Second, once the user selects a theme, the system suggests multiple "Socratic questions" [18], an evidence-based approach that encourages critical thinking and self-reflection. These questions prompt users to explore underlying emotions and perspectives that may not have surfaced in their initial narrative.

***DR2. Provide guidance to support expression without directing content.*** In our formative study, professionals highlighted the usefulness of providing subtle guidance to help individuals articulate their thoughts and emotions. They emphasized that people often struggle with open-ended or thought-provoking questions, which can be cognitively overwhelming. In the context of writing for well-being, the benefit lies in the individual's process of translating their experiences into language [89]. To preserve this, our system avoids directly shaping the content of users' expressions (*e.g.* generating direct sentences for users to adopt [48]), instead offering scaffolding that gently supports the reflective process. To achieve this balance, we introduced features like "Keywords" and "Comments." The keyword list offers users a subtle prompt by suggesting relevant concepts or terms that might help them think through a particular Socratic question, while still leaving room for their own interpretation. Additionally, the system generates comments that provide hints or feedback without making direct assertions about the user's narrative.

***DR3. Provide a tangible summary of the progress.*** In our formative study, professionals emphasized the importance of restating or summarizing client's thoughts back to them, as this helps foster reflection by offering an objective perspective on the client's experiences. They noted that while the summary might be subjective from the therapist's viewpoint, it allows individuals to compare their internal thoughts with an external interpretation, serving as a valuable tool for stepping back and critically evaluating their journey. In line with this practice, we introduced the "AI Summary" feature, designed to offer users an objective summary of their reflections, distilled from the user's themes, questions, and answers, reflecting the key points of their exploration.

## 4.2 System Design and User Interface

EXPLORESELF is designed to assist users in navigating, exploring, and making sense of personal challenges through writing with structured and adaptive guidance by an LLM. The application consists of three primary phases of user engagement: (1) *Writing the initial narrative* (Figure 2), (2) *Exploring the narrative* (Figure 3),

**Figure 2: The Initial Narrative page of EXPLORESELF. Writing the initial narrative ⓐ is a starting point for providing the basis clues to the system. Clicking the 'Start Exploration' button ⓑ leads to the Exploration page (Figure 3-A).**

and (3) *Wrapping up and summarization* (Figure 4). Each phase was carefully designed to encourage different aspects of self-reflection on and understanding of the personal challenges, helping users process their thoughts and emotions effectively. In the following, we describe each phase and interactions through a usage scenario: *Jane is a retiree who looked forward to enjoying the freedom she earned after decades of hard work. She had plans to explore new hobbies and focus on self-development. However, her daughter asked Jane to care for her newborn during the daytime. Although Jane loves spending time with her grandson, she sometimes feels that her own life has become overshadowed by her new responsibilities. The freedom she anticipated seems out of reach, and she occasionally struggles with a sense of purposelessness, yearning for time to pursue her personal growth. Now that she turns to EXPLORESELF to process this challenge.*

*4.2.1   Writing the Initial Narrative.* Jane opens EXPLORESELF and the web page shows the *Initial Narrative* page (Figure 2), which prompts her to articulate her challenges. The interface presents a borderless text area encouraging Jane to write freely about what weighs on her mind and her thoughts and feelings associated with it (ⓐ in Figure 2). Jane jots down her story and clicks the 'Start Exploration' button (ⓑ in Figure 2), which sends her to the Exploration phase, where the system facilitates a more structured process of introspection and analysis.
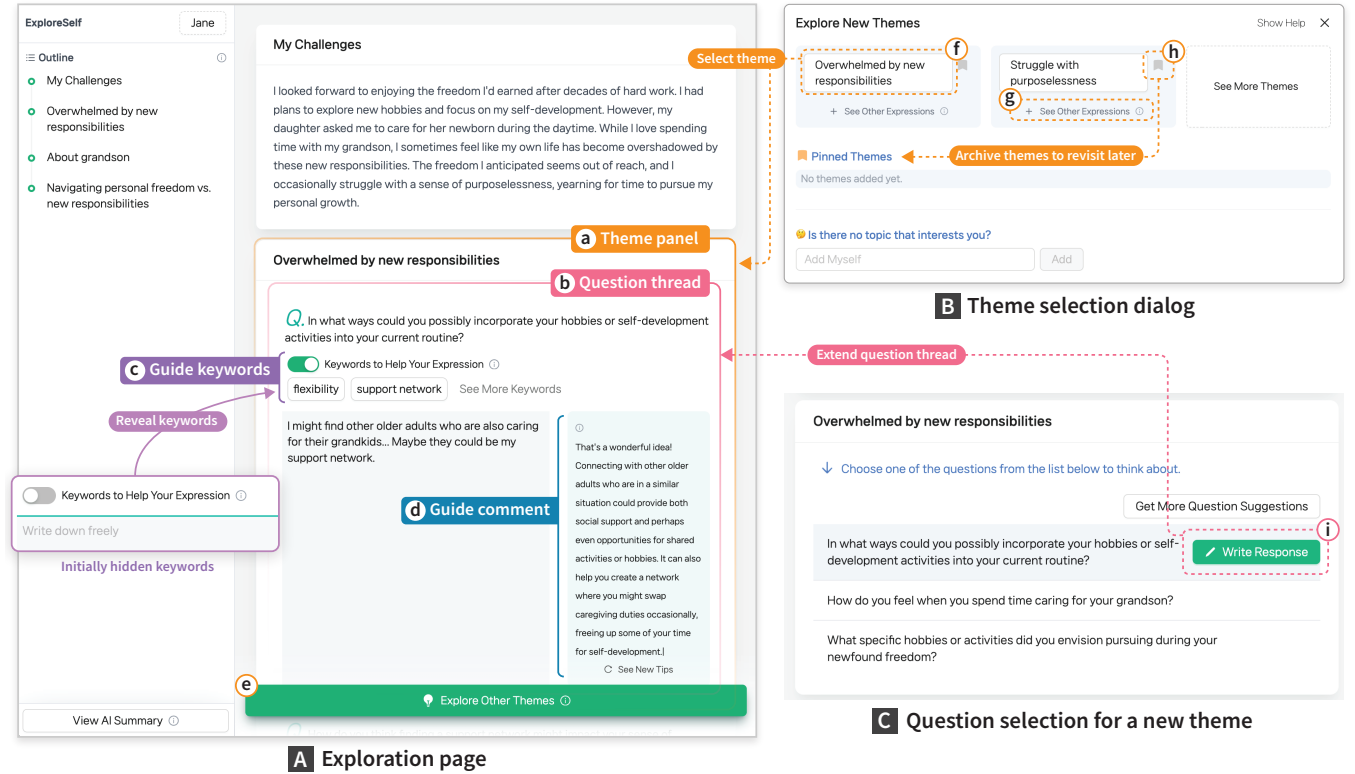
*4.2.2   Exploring the Narrative.* The **Exploration** page (Figure 3-A) allows Jane to explore and reflect on her challenges by navigating both the breadth and depth of her narrative.

*Themes: Expanding breadth.* Jane clicks the 'Explore Themes' button (ⓔ in Figure 3-A), and it opens the **Theme Selection** modal dialog (Figure 3-B). The dialog lists two AI-generated themes related to her initial narrative: 'Overwhelmed by new responsibilities' and 'Struggle with purposelessness.' She would like to explore both themes but saves the second one for later by clicking the bookmark icon (ⓗ in Figure 3-B), which stores the theme in **Pinned Themes**. She then clicks the 'Overwhelmed by new responsibilities' theme (ⓕ in Figure 3-B). And a new **Theme panel** (Figure 3-C) appears on the main page.

*Question Threads: Expanding depth.* The theme panel lists three AI-generated Socratic questions that may encourage deeper reflection on the current theme. Of these questions Jane selects the first one: "In what ways could you possibly incorporate your hobbies or self-development activities into your current routine?" (ⓘ in Figure 3-C) creating a new **Question panel** (ⓑ in Figure 3-A) on the theme panel. The question panel contains a text field to write down answers. Next to it, an **AI comment** (ⓓ in Figure 3-A) is generated, encouraging Jane to think about small pockets of time. To get more guidance for writing answers, Jane turns on the switch for the **Keyword list** (ⓒ in Figure 3-A) above the text field, which would reveal two relevant keywords, 'flexibility' and 'support network.' Inspired by the support network keyword, Jane jots down the idea of finding peer older adults who also care for grandkids to form a support network. She clicks the 'See New Comment' button (ⓓ in Figure 3-A) to refresh the comment, and the new comment acknowledges her answer. Jane goes on this thread of questions; she clicks the 'Get More Question Suggestions' button (Figure 3-C) on the current question panel, and the Theme panel generates three new questions based on her question and answer. She selects 'How do you think finding a support network might impact your sense of overwhelming responsibility?' that directly follows up her previous thread. She continues exploring the theme by creating and answering AI-driven questions.

Satisfied with the exploration of the ideas of the support network, Jane turns to a new theme. She again clicks the 'Explore Other Themes' button, and the Theme selection dialog suggests new themes generated considering her activities so far. However, no suggested themes inspire her. So Jane **creates her own theme**, "About grandson," by typing it on a text field on the Theme selector dialog, to reflect on her grandson. The new Theme panel suggests questions related to her custom theme, and Jane selects the question, "What are the aspects of spending time with your grandson that bring you joy and fulfillment?" Upon this question, her perspectives on caring for her grandson transitioned from being overshadowed to being fulfilled. She continues exploring through her challenging story by adding themes and questions. She also goes back to see past themes and her writing by clicking themes on the **Overview** list (Figure 3-A, top-left). After a while of exploration, Jane feels that her exploration is saturated. She clicks the 'View AI Summary' button (Figure 3-A, bottom-left) to wrap up the exploration.

*4.2.3   Wrapping Up and Summarization.* Jane enters the **AI Summary** screen (Figure 4). On the left column, she can review a comprehensive list of themes, questions, and answers (ⓐ in Figure 4).

**Figure 3: Detailed interactions of the Exploration page [A] of ExploreSelf. The main scroll panel vertically appends a theme panel ⓐ when the user creates a new theme ⓕ on the Theme selection dialogue [B]. The theme panel initially displays only the list of AI-suggested questions regarding the theme [C], and the user can populate the question thread ⓑ by selecting the next question they want to answer ⓘ. While answering a question, users can reveal keywords and request more ⓒ. The system also provides a guide comment ⓓ, which can be regenerated upon request.**

On the right column, an AI-generated summary is displayed (ⓑ in Figure 4). The summary distills the key themes and insights from the user's text into a concise essay. Using the back button (ⓓ in Figure 4), she can also go back to the Exploration screen and return with updated exploration paths. Jane reads through the AI summary and ends the exploration.
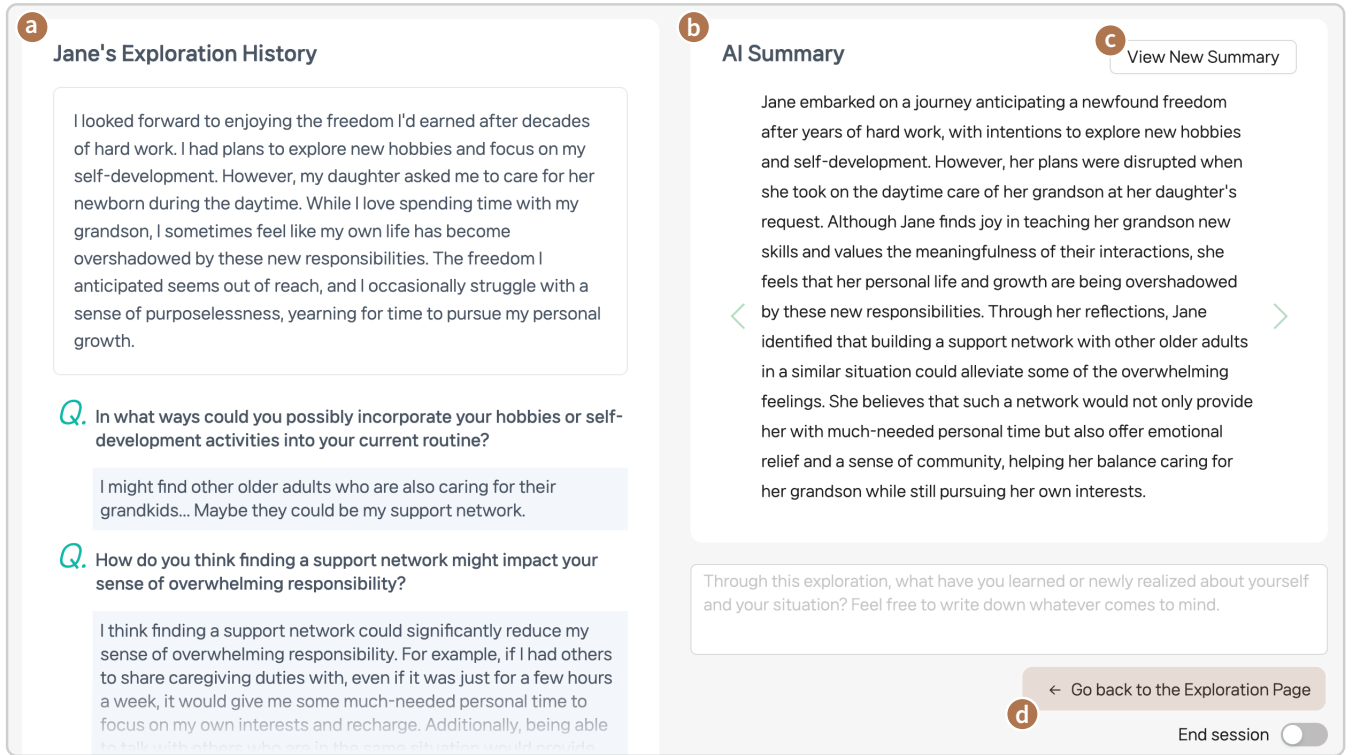
## 4.3 Generative Pipelines

In this section, we describe the generative pipelines incorporated in ExploreSelf to support the generation of themes, questions, keywords, comments, and summaries. Figure 5 illustrates the data flow and a gist of LLM instructions for each pipeline. The pipelines refer to the current exploration status (see Figure 5, left), including the initial narrative, themes, questions and answers, and keywords and comments selectively depending on the type of pipeline. The exploration status is inserted to the structured input instruction formatted in XML. The underlying LLMs run on instructions following the chain-of-thought prompting approach [120], where the model is instructed to provide meta-output (see thin texts in Figure 5, right), such as rationales of the output, along with the requested output to enhance the reliability of generation. We consistently applied a

'therapeutic assistant' persona to the instructions (see dark boxes in Figure 5, right) to better contextualize the generation [33, 119].

The **Theme Generation** step (Figure 5-[A]) identifies themes that arise from their personal narratives and previous interactions with the system. The model is instructed to name themes that closely align with the user's language and expressions. In case the user does not grasp the meaning of the original theme name, the model also provides *alternative expressions* to allow the user reveal them on the Theme selector dialog (see ⓖ in Figure 3-[B]). To further guide this behavior, the model clarifies which part of the initial narrative each theme stems from (see the white box in Figure 5-[A]) (*c.f.*, Section B.1). The **Question Generation** step (Figure 5-[B]) suggests probing questions for a given theme. The model is instructed to craft socratic questions [18] that encourage users to navigate deeper into their thoughts and feelings. Each question is designed to resonate with the user's language and context (*c.f.*, Section B.2).

To further guide the user while reflecting on the questions, the system also provides **keywords** (Figure 5-[C]) and **comments** (Figure 5-[D]). When generating keywords, the model is instructed that they should (1) serve as cognitive scaffolding, (2) activate potential blind spots for the user, and (3) be relevant to the user's background and core values (*c.f.*, Section B.3). As for the comments, the model

**Figure 4: The AI Summary page of ExploreSelf where the users can overview their exploration history (a) and the AI-generated summarization (b). The 'View New Summary' button (c) generates a new summary. Users can go back to the Exploration page (Figure 3-A) using the 'Go back' button (d).**

is instructed to first decide the category of the comment from one of the predefined categories: (1) tips for approaching the question, (2) encouraging feedback, (3) sub-questions elaborated from the questions, (4) insightful comment, and (5) others (See Section B.4).

When generating **AI summary** (Figure 5-E), the pipeline distills the user's themes, questions, and answers into an essay in several paragraphs. The model is instructed to generate a concise and focused story, using the user's own language and expressions where appropriate. The summary is intended to highlight key themes, emotions, and notable progress, offering a clear snapshot of the user's pathway of exploration. Also, the model tries to make the summary commensurate with the volume of the content of the user's writing, not extrapolating details and staying grounded in the user's actual input. In particular, the model prompt (*c.f.*, Section B.5) includes guidelines "Use the user's own language and expressions where appropriate" and "Keep the summary realistic, proportional to the content of the user's log, and grounded to user's input."
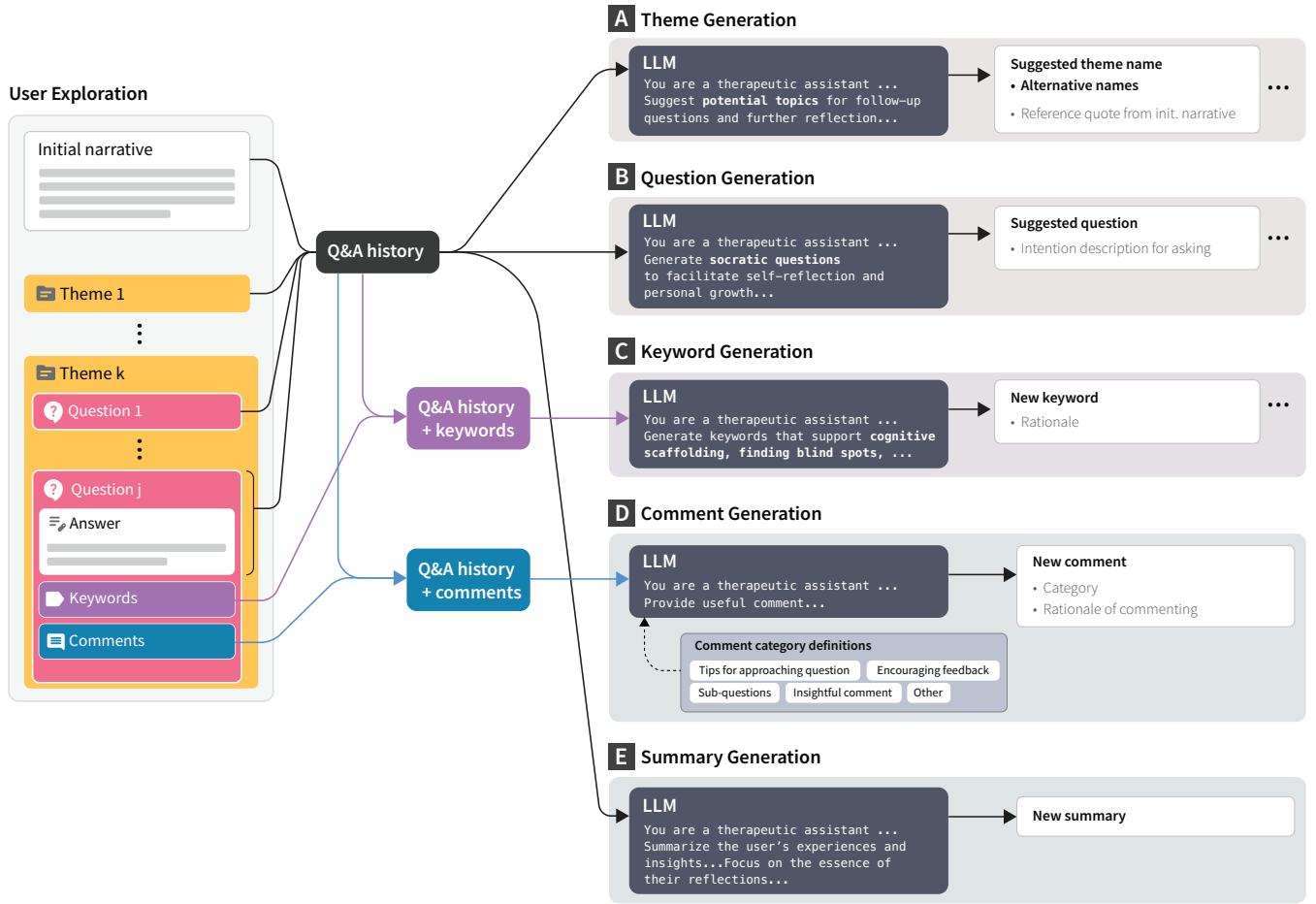
### 4.4 Implementation

ExploreSelf consists of (1) a backend server that manages user data and LLM pipelines and (2) a frontend web interface, both written in TypeScript [70]. The server runs on the Express.js [30] framework, providing core functionalities via REST APIs and storing the user data in MongoDB [73]. We implemented the LLM pipelines using LangChain.js [53] incorporating OpenAI [80]'s GPT-4 ChatCompletion API. As LLM's capabilities of dealing with non-English texts are inferior to equivalent English texts due to the inefficient tokenization [92], we chose gpt-4o[1] as an underlying model, which involved an improved Korean tokenizer with state-of-the-art performance. The web interface was implemented using React [69].

## 5 User Study

We conducted an exploratory study to examine how individuals interact with ExploreSelf and how it may foster exploration and reflection on personal challenges. As part of the study, participants explored their own challenges using ExploreSelf on their computers. The study was conducted online via Zoom to allow participants to engage from a comfortable and secure environment of their choice. This online approach enabled participants to interact with ExploreSelf at their own pace and in their preferred settings, which has been shown to foster more authentic and reflective responses [38, 104]. We refined both the system and the study protocol through three pilot sessions with Korean adults. For example, we incorporated the Pinned Themes feature, enabling users to pin and revisit previously generated themes. We also revised the wording used in the interface and the tutorial procedure to clarify that the system is not intended to provide solutions or counseling but to

---

Figure 5: The overview of pipelines to generate themes A, questions B, keywords C, comments D, and the summary of the exploration history E in the middle of exploration. Each generative pipeline incorporates an LLM inference that receives the current exploration history starting from the initial narrative as an input. The LLM inference is driven by chain-of-thought style instructions, providing descriptions and rationales for each output to enhance reliability.

support thematic exploration. The study protocol was approved by the public institutional review board of the Ministry of Health and Welfare of South Korea.

## 5.1 Participants

We recruited 19 participants (P1–19; 10 female) by advertising the study on a local social community platform and online university communities in South Korea. Our inclusion criteria were adults who (1) are 19 years old or older; (2) are native Korean speakers; (3) are currently staying in Korea; (4) can join an online study in a private and comfortable place; and (5) are motivated to better understand their personal challenges. In particular, we ensured participants met criteria (2) and (3), (2) for them to be able to express their personal challenges in the most comfortable language as EXPLORESELF was implemented in Korean. (3) was to make sure the participants were in an accessible proximity, allowing them to engage in the study from a private space of their choosing, where they could write without researchers physically present while still

being accessible for the study. Please refer to Section 5.2 for more detail. The advertisement included a link to a screening survey that asked basic demographic information (*i.e.*, age, gender, educational level, employment details, as well as background questions about prior experience with AI and personal endeavor to process personal challenges, if any. We did not screen participants for the presence of mental illness, as EXPLORESELF was designed to support the general population's mental wellbeing rather than serve as a psychotherapeutic intervention.

Table 2 shows the demographic composition of our study participants. Participants were from 22 to 65 years old (*avg* = 37.53). The participants comprised 11 full-time or part-time workers in various sectors, including distribution and elderly welfare, three self-employed local business owners, and the rest consisting of two graduate students, an undergraduate student, an unemployed person, and a homemaker. Fifteen out of 19 participants had tried conversational AI assistants like OpenAI's ChatGPT and Google Bard (now Gemini) more than once in their daily lives, and six of

**Table 2: Demographic information of the participants and their self-report frequency of using AI assistants like ChatGPT [79] and Bard [32].**

| Alias | Age | Gender | Occupation | Frequency of AI assistant use |
|---|---|---|---|---|
| P1 | 39 | Female | Full-time (IT) | At least once a day |
| P2 | 28 | Female | Graduate student | Within 3 times a week |
| P3 | 28 | Male | Unemployed | At least once a day |
| P4 | 44 | Male | Self-employed | Within 3 times a week |
| P5 | 26 | Male | Graduate student | At least once a day |
| P6 | 44 | Female | Self-employed | At least once a day |
| P7 | 25 | Male | Homemaker | Never |
| P8 | 53 | Female | Part-time (Sales) | Tried once or twice |
| P9 | 27 | Male | Self-employed | 4-6 times a week |
| P10 | 41 | Female | Full-time (Energy) | Never |
| P11 | 36 | Male | Full-time (Finance) | Tried once or twice |
| P12 | 22 | Male | Undergraduate student | At least once a day |
| P13 | 37 | Female | Preschool teacher | Tried once or twice |
| P14 | 65 | Female | Part-time (Elderly welfare) | At least once a day |
| P15 | 40 | Female | Adjunct instructor | 4-6 times a week |
| P16 | 32 | Female | Full-time (Healthcare) | Tried once or twice |
| P17 | 39 | Male | Private academy instructor | 4-6 times a week |
| P18 | 38 | Female | University lecturer | Within 3 times a week |
| P19 | 49 | Male | Full-time (Distribution) | Never |

them used AI assistants for general purposes at least once a day. Regarding the efforts to process personal challenges, a majority of participants (13 out of 19; 68%) had explored books or materials like YouTube videos about mental health, seven participants had had professional counseling, and five participants had tried AI chatbots like ChatGPT and Bard to discuss their challenges. To compensate for their participation, we offered a 50,000 KRW (approx. 38 USD) online gift card.

## 5.2 Safety Considerations

Although our target participant group did not specifically include individuals experiencing mental illness, we prepared safety protocols for the study based on the work of O'Leary et al. [78], which involved peer support chat amongst individuals experiencing mental illness. This was a mindful procedure considering that ExploreSelf encourages participants to disclose personal issues that might be sensitive and potentially traumatic, and participants could have mental health issues that were not disclosed to the experimenter.

Prior to the study, participants were provided with a list of emergency contact numbers and municipal support resources to ensure they had access to immediate help if needed. During the study, participants were encouraged to take breaks at any time for any reason, upon notifying the researcher through a real-time messaging system if they required one. The first author carefully monitored

participant interactions with ExploreSelf to identify any potential risks, particularly focusing on signs such as mentions of self-harm or harm to others, as well as indications of increased distress, frustration, or exposure to triggering LLM-generated content. If any of these occurred, the first author would promptly engage with the participant to address the issue and take appropriate actions. These actions include immediately halting the session to secure participant safety; contacting municipal support resources on behalf of the participant in case of requiring help; and referring the participant to the healthcare center within the affiliated institution that included a consulting psychiatrist, should the participant needed to consult a clinician at once. We informed the participants of our partial compensation policy in case of any emergencies. In the course of the study progression with all 19 participants, no such event occurred.

## 5.3 Study Setup and Procedure

Each participant was invited to a 90-minute study session remotely via Zoom video call on their computer. Instructions for installing and using the remote sharing software were prepared in case of need. One researcher administered the sessions as an experimenter. The study session consisted of four parts: (1) briefing, (2) tutorial, (3) exploration with ExploreSelf, and (4) debriefing.

***Briefing.*** The experimenter shared her screen with presentation slides (refer to our supplementary material) and described the motivation and goal of our study. To encourage participants to authentically engage with the system, we informed them that the goal of this study is not to evaluate or address their specific personal challenges but to understand how our tool helps their exploration and reflection on those challenges.

***Tutorial.*** After explaining the goal of the study, the experimenter gave a 10-minute tutorial covering the main features of ExploreSelf using a slide presentation. We masked out any dynamically generated or user-provided texts from the screenshots to minimize the potential bias from example contents that might influence participants' preconceptions of personal challenges and writing.

***Exploration.*** After the tutorial, the experimenter shared a link to access ExploreSelf and invited the participant to a Zoom breakout room, a virtual space where the participant could stay alone and focus. The experimenter was able to monitor the real-time interaction logs on ExploreSelf to observe the activity patterns. The experimenter and the participant were connected via an instant messenger to communicate about any issues the participant may have during the exploration including help requests. The participant then freely used ExploreSelf to explore their challenge. There was no minimum time requirement, and participants could finish the session whenever they felt they had sufficiently used it, for up to 45 minutes. At the 45-minute mark, the researcher checked in with participants if they needed more time. If so, they were allowed up to an additional 15 minutes, with a maximum session time of 1 hour. Since participants did not have free practice in the tutorial phase, the system offered on-demand help popups on major interface components such as the Theme Selector dialog. To assess the impact of the exploration on participants' attitudes towards their challenges, we asked participants to rate their perceived agency by completing the validated Korean version [16] of the Pathways Subscale of the

State Hope Scale [112], before and after of this phase. This 4-item measure asks participants to rate statements such as "*There are lots of ways around any problem that I am facing now*" on an 8-point scale, ranging from 1 (Definitely false) to 8 (Definitely true). The Pathways Subscale of the State Hope Scale (refer to Appendix C), used in this study, captures individuals' perceptions of their ability to achieve their goals [13, 103, 112]. Scores range from 4 to 32, with higher scores reflecting a stronger belief in the availability of pathways to their goals.

***Debriefing.*** At the end of the session, we removed the breakout room and conducted a semi-structured interview for about 30 minutes. Before starting the interviews, we confirmed their willingness to proceed. We asked participants about their experiences with EXPLORESELF, their rationales for adding and moving between themes and questions, receptiveness to AI-driven guidance, and how the interface helped or hindered their exploration and reflection.

## 5.4 Data Analysis

We employed both quantitative and qualitative methods to analyze our data, in order to gain a comprehensive understanding of participant engagement and interaction with EXPLORESELF. First, to understand what types of personal challenges participants would explore on our system, participants' initial narratives were categorized through an iterative open-ended coding process conducted by two authors. The first author, who has two years of experience in HCI and mental health research and administered the user study, first open-coded the initial narratives. As the initial narratives usually described multiple aspects, we multi-coded the narratives. Another researcher with a PhD in computer science and 12 years of experience in HCI and health informatics research participated in multiple rounds of discussion to finalize a codebook for sub-category labels. The full research team then iterated multiple sessions of discussions to group the subcategories into five broader categories. Although the coding was primarily bottom-up and grounded from the data, some category labels were informed by Sharma *et al.* [106]'s definitions of personal issues (*e.g.*, *parenting*, *trauma*) and the life values set used in Acceptance and Commitment Therapy [36] (*e.g.*, *social relationships*). The categorization is reported in Table 3.

To analyze their overall usage and interaction patterns, We collected interaction logs to track key behaviors, such as the number of themes and questions selected by participants, the number of interactions with keywords and comments, and phase-switching behaviors (*e.g.* between exploration and overview phase). These logs were analyzed to compute descriptive statistics, including the minimum, maximum, mean, and standard deviation for user engagement across different system features.

The Zoom sessions for debriefing interviews were recorded and transcribed later. We qualitatively analyzed the debriefing transcripts using thematic analysis [11] to identify key themes regarding participants' experiences with EXPLORESELF. The first author open-coded the transcripts, going through several rounds of iterations. The full research team then discussed and identified patterns and themes through multiple rounds of peer debriefing. From this coding, we surfaced the key themes around the participants' sense of agency while using EXPLORESELF and how the system features supported the exploration, reflection, and self-expression.

## 6 Results

In this section, overall system usage based on quantitative metrics from interaction logs and surveys. We then cover how the system influenced participants' sense of agency during the exploration. Drawing on the qualitative analysis of the debriefing, we describe how themes and questions supported the user-directed navigation of personal challenges, and how keywords and comments helped participants' self-expression. Lastly, we report participants' perceived utility of EXPLORESELF.

## 6.1 Overall System Usage

Participants wrote about a wide range of personal challenges in their initial narratives, from personal relationships to dilemmas in life decision-making. Table 3 summarizes the categories of initial narratives participants entered. Participants mostly engaged with the challenges regarding *personal relationships* (12 participants; 63%), followed by *self-development* (9 participants; 47%), *self-identity & personal character* (8 participants; 42%), *stability in life* (6 participants; 32%), and *health* (5 participants; 26%). The initial narratives often described multiple interconnected topics.

**Table 3: The categorization of participants' initial narratives. The sub-categories are multi-coded to the initial narratives and grouped into categories. So, the number of participants for each category exceeds the total number of participants.**

| Categories | Description of categories | Sub-categories |
|---|---|---|
| **Relationships** (12 participants) | Personal interactions and societal influences affecting connections with others. | Family conflicts[P1, P8, P14, P15], parenting[P1, P7], social relationships[P6, P8], romantic relationships[P3, P16], societal stigma[P4], societal pressure[P1, P11, P13, P19] |
| **Self-development** (9 participants) | Challenges in personal growth, career, and managing goals. | Academic stress[P5], job dissatisfaction[P10], time-management concern, goal fulfillment, career uncertainties[P1, P5, P6, P9, P12, P14, P15, P17] |
| **Self-identity & character** (8 participants) | Issues related to personal identity, self-perception, and existential concerns. | Aging[P8, P14], sexual identity[P3], existential concern[P8, P16], self-esteem issue[P2, P7, P8, P13, P19] |
| **Stability in life** (6 participants) | Challenges affecting financial and personal stability. | Financial crisis[P8, P9, P11], financial instability[P5, P9, P18], life transitions[P16] |
| **Health** (5 participants) | Mental, emotional, and physical well-being challenges. | Managing anxiety[P16, P18], trauma[P4, P8], emotional struggles[P2], physical health concerns[P8] |

**Table 4: Summary of syllable counts of the participants' inputs and the number of exploration and guide entities by participant. Note that for comments we counted only the comments that were manually requested by participants.**

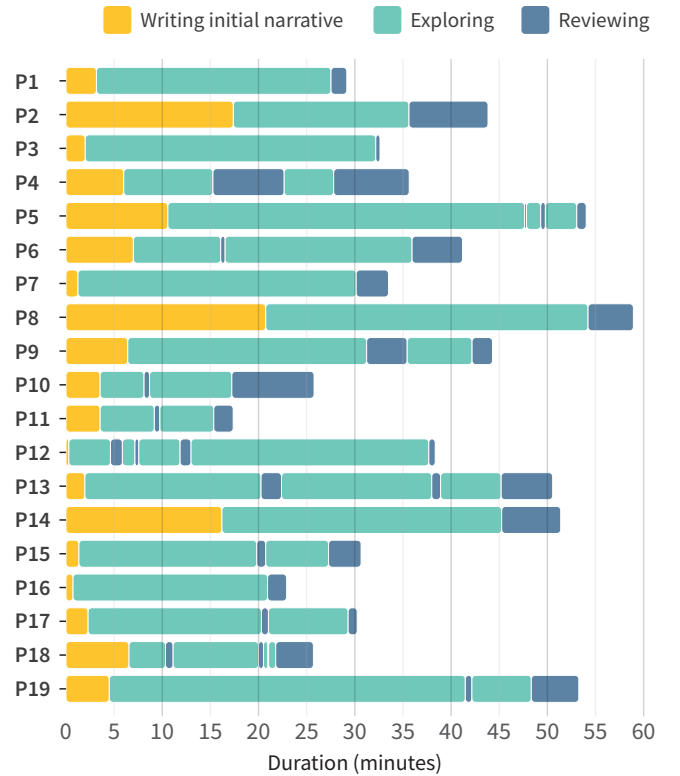| Metric | Mean | SD | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 | P11 | P12 | P13 | P14 | P15 | P16 | P17 | P18 | P19 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Syllables of initial narrative | 233.63 | 299.98 | 233 | 478 | 39 | 215 | 313 | 445 | 42 | 1347 | 100 | 180 | 198 | 13 | 137 | 1,534 | 25 | 34 | 111 | 239 | 127 |
| Total syllables of responses | 745.47 | 349.65 | 1,261 | 568 | 699 | 388 | 1,273 | 1,019 | 541 | 1,278 | 571 | 325 | 381 | 496 | 1,381 | 417 | 824 | 508 | 864 | 465 | 905 |
| # of themes | 4.89 | 2.26 | 7 | 3 | 2 | 2 | 9 | 5 | 5 | 5 | 6 | 4 | 3 | 4 | 5 | 6 | 4 | 11 | 3 | 4 | 5 |
| Total # of questions | 11.47 | 7.28 | 17 | 7 | 18 | 3 | 11 | 4 | 15 | 6 | 5 | 4 | 8 | 28 | 12 | 9 | 10 | 12 | 22 | 4 | 23 |
| Total # of revealed keywords | 11.79 | 8.69 | 0 | 9 | 0 | 27 | 6 | 11 | 29 | 18 | 8 | 9 | 24 | 15 | 4 | 18 | 20 | 4 | 6 | 9 | 7 |
| Total # of comment requests | 6.47 | 8.26 | 0 | 1 | 0 | 2 | 0 | 5 | 18 | 4 | 1 | 0 | 6 | 1 | 11 | 3 | 10 | 27 | 9 | 1 | 24 |

Table 4 summarizes the syllable counts of participants' written inputs (*i.e.*, initial narratives, and responses to questions) and the number of themes, questions, keywords, and comments generated or requested by participants. The average syllable count of the initial narratives was 233.63 with high variance ($SD = 299.96$, $min = 13$ [P12], $max = 1347$ [P8]). The total syllable count of all responses per participant was 745.47 on average ($SD = 349.65$, $min = 325$ [P10], $max = 1381$ [P13]). There was a moderate correlation between the syllable count of the initial narrative and the total syllable count of the responses, with a Pearson correlation coefficient of 0.49 ($p = 0.03$), indicating a statistically significant relationship.

There was also a high individual variance in how participants spent time across screens. Figure 6 illustrates the participants' timelines of the exploration phase, split by the type of screens they stayed on. None of the participants took breaks during their engagement, and they took an average of 37.86 minutes in the entire exploration phase ($SD = 11.58$, $min = 17.37$ [P11], $max = 58.93$ [P8]). While writing the initial narrative, participants spent an average of 5.82 minutes ($SD = 5.40$, $min = 0.3$ [P12], $max = 17.38$ [P2]; see yellow bars in Figure 6). Then they spent an average of 31.77 minutes ($SD = 9.42$, $min = 13.82$ [P11], $max = 48.8$ [P13]) for both exploring and reviewing the AI summaries (see green and blue bars in Figure 6).
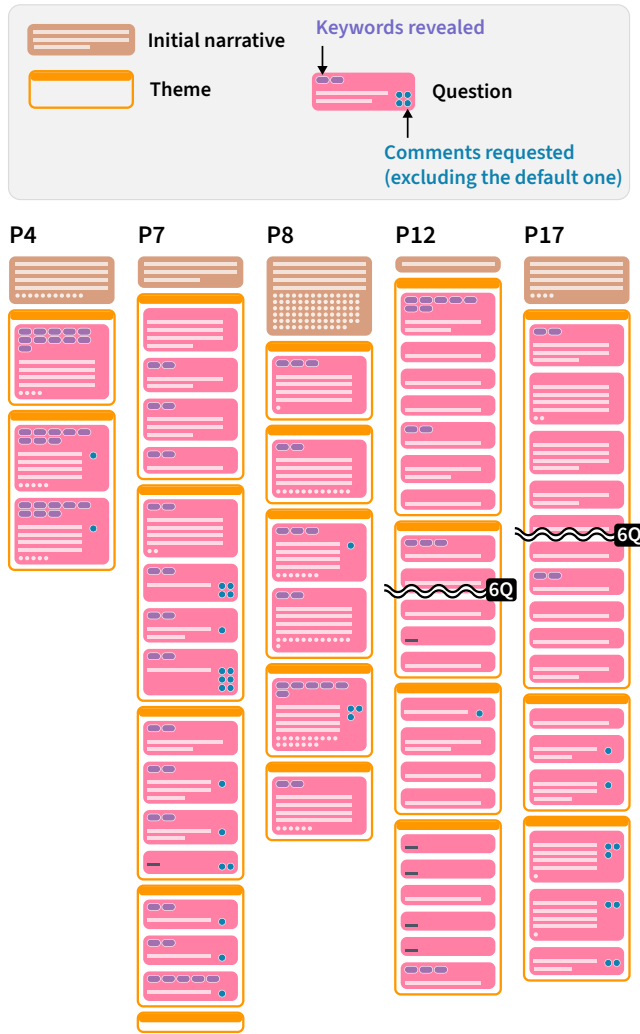
Participants showed diverse patterns in exploration, engaging with a varying number of themes and questions suggested by ExploreSelf. Table 4 reports the descriptive statistics of participant engagement. On average, participants engaged with 4.89 themes ($SD = 2.26$, $min = 2$ [P3, P4], $max = 11$ [P16]) and 11.47 questions ($SD = 7.28$, $min = 3$ [P4], $max = 28$ [P12]). The number of questions per theme averaged 2.58 ($SD = 2.61$, $min = 1$ [P4, P5, P6, P8, P9, P10, P11, P13, P14, P16, P18], $max = 16$ [P17]). The total syllable count of responses written during the exploration phase averaged 796.95 ($SD = 464.44$, $min = 269$ [P12], $max = 1674$ [P13]). Figure 7 illustrates the results of exploration captured from selected participants, who showed distinct engagement patterns with different components. For example, P4 actively leveraged keywords (purple capsules in Figure 7), and P7 generally used both the keywords and comments (blue dots in Figure 7) throughout all questions. P8 wrote an extreme amount of text for the initial narratives and answers, and P12 showed a tendency to write a short narrative and answers. Lastly, P17 dove into the first theme with many questions.

## 6.2 Sense of Agency over Exploration

Participants' perceived agency on their challenge increased after using ExploreSelf: The pathway subscale score, which we collected before and after the system use, significantly increased from 22.32 ($SD = 4.91$) to 24.95 ($SD = 5.86$) by 2.63 on average ($t(18) = 2.80$, $p = 0.012^*$). Cohen's $d$ of the two means was 0.66, indicating the



**Figure 6: The timelines of participants during the exploration phase. The X-axis indicates elapsed minutes. The yellow bars denote the periods spent writing the initial narratives in the Initial Narrative page (Figure 2), the cyan bars denote the periods staying in the Exploration page (Figure 3-A), and the blue bars denote the periods staying in the AI Summary page (Figure 4). Note that participants were able to move between the Exploration and the AI Summary pages.**

**Figure 7: The results of the exploration of selected participants with distinct patterns. The brown boxes denote the initial narratives, the yellow frames denote themes, and the pink boxes denote questions. Number of keywords that participants revealed is depicted as purple capsules. The blue dots denote the comments that participants additionally requested. The white lines and dots illustrate the amount of text for the initial narratives and answers (the lines and dots both denote 15 syllables.)**

effect size of the gaps falling between medium (0.5) and large (0.8). Specifically, the scores of 15 out of 19 (79%) participants increased.

In debriefing, participants expressed varying perceptions of their agency while using ExploreSelf, with several factors influencing their sense of control. Some participants felt that the **acknowledgment of their wordings in the AI-generated text** played a crucial role in shaping the experience. P11 observed that while the AI sometimes paraphrased his inputs into different expressions in its generation, this paraphrasing gave him the impression that the AI was taking an active role. However, he still felt a sense of

agency because the AI-generated texts were grounded in his original wording, highlighting that his expressions ultimately drove the interaction. P8 echoed this sentiment, noting that the system's ability to suggest themes based on her input made her feel that she was leading the process.

Participants also mentioned that the **generation latency** of AI also influenced their sense of control. For instance, P14 highlighted that while she retained some control, *"the AI's faster processing of information made it feel more like the system was taking the lead at times, especially when [her own] thoughts formed more slowly."*

The **maturity of reflections** further enhanced participants' perceived control over the process. P13 described how, in the early stages of each theme, she felt more guided by the AI, but as she began to clarify her thoughts, she began to feel more confident and in control in the theme thread, feeling that she was leading the exploration. This growing clarity allowed P13 to *"confidently decide when to explore a theme more deeply and when to move on to a new one"* when she *"felt [she] had fully explored the current one."* This gradual shift from AI-driven to self-directed reflection was echoed by a few more participants who appreciated the system's guidance in the initial phases but ultimately took ownership of the process as their reflections unfolded.

On the contrary, participants felt that their autonomy was compromised by the AI's influence, especially when the AI's suggestions of themes and questions derailed from their primary concerns and steered them toward unintended directions. P4 remarked, *"The problem I originally had was related to my family, but the AI kept focusing on my band activities as if there were an issue there. Even though I knew that wasn't the case, I ended up revisiting the topic, which made the process take longer."*

In general, participants appreciated the way ExploreSelf visualized and dynamically structured the breakdowns of their thoughts into a coherent landscape. Ten participants highlighted how this experience differed from their self-journaling practices. P18, for instance, shared that when reflecting on her own, her thoughts would often *"spiral in a negative direction"* with no clear way to shift perspectives. In contrast, the system allowed her to *"step back"* and revisit previous thoughts easily, helping her avoid getting stuck in a mental loop. Similarly, P7 remarked that if she had tried to write about her stress and various life challenges on her own, her thoughts would have *"jumped from one issue to another,"* leading to a scattered reflection. However, with the AI's structure, she could *"stick to one theme"* at a time, making the process more focused and less overwhelming.

## 6.3 User-directed Navigation

Participants navigated through themes and questions by making deliberate choices about whether to explore deeper into a particular theme (*i.e.*, add a new question to the thread) or move on to a new theme. Based on the findings from debriefing interviews, we present how participants interacted with the themes and questions in the pathways of processing their thoughts and emotions.

*6.3.1 Expanding Viewpoints through Themes.* Participants generally found the LLM-generated themes were reflective and responsive to their previous input, with many noting the relevance and accuracy of the suggestions. They reported that they selected the

themes by prioritizing clarity and personal relevance, aiming to focus on issues that they felt were actionable and meaningful. For instance, P18 preferred more *"condensed"* words as themes, believing they would lead to more insightful AI-generated questions. Others, like P5 and P8, tended to select themes that aligned with their immediate concerns, allowing them to focus on issues they could actively address. At the same time, however, participants avoided themes that felt overly literal or much echoed their initial expression. For example, P2 preferred themes that *"offered some interpretation or deeper insight,"* because she wanted to *"explore the heart of the matter, rather than focusing on surface-level concerns."* This balance between relevance, clarity, and depth helped participants feel more in control of their reflective process.

Participants also emphasized that the suggestions of alternative wordings for the same theme allowed them to engage in their narrative from multiple angles, broadening their perspectives and deepening their reflections. This functionality encouraged participants to think about the same theme in different ways, as seen when P14 explored the theme of 'retirement' from multiple viewpoints, such as " *anxiety about retirement,*" " *purpose and direction of life after retirement,*" and " *feeling unprepared for old age."* These varied expressions also provided flexibility in how participants responded to the themes, prompting different sets of questions depending on how a theme was framed, and also expanded their capacity for self-expression. As P9 noted, the system's varied prompts allowed him to *"respond more creatively and diversely"* by offering new ways to articulate their thoughts.

Participants tended to move on to a new theme when they felt they had gained enough insight or a good understanding of a particular issue. For example, P13 shifted her focus toward solutions once she understood that her low self-esteem was the root of her challenge: " *I decided to think about how to address it, rather than keep going deeper into why my self-esteem was low."*

*6.3.2 Organizing Angles and Attitudes through Questions.* Participants found that being presented with multiple questions helped them consider their challenges from various angles, exploring different aspects and depths of their situation. P7 noted how it allowed him to see *"how different questions can expand one thought into multiple pathways,"* likening the process to a *"mind-storming"* session where diverse questions spurred deeper thought. P11 added that seeing questions tied to his previous responses helped *"take the reflection process step by step, encouraging [him] to explore aspects they may not have considered otherwise."* Some participants viewed the act of deciding which question to engage with based on what seemed most effective at the moment as a reflective practice in and of itself. P5 mentioned, *"If new challenges arise, I can think through them using these types of questions, such as considering why the problem occurred, what abilities I have to resolve it, and how I might approach the solution."*

The provision of multiple follow-up questions enabled participants to steer the reflection process, often balancing their cognitive and emotional discomfort with the question's significance regarding their personal concerns. Many chose questions that carried the most weight on their issues, even if those questions were mentally burdensome to answer. As P18 shared, she prioritized *"the most*

*pressing concern,"* recognizing that addressing these difficult topics would ultimately help her confront more difficult issues.

Participants also valued questions that challenged their usual thinking patterns, prompting reflections on aspects they had never considered before. P2 noted, *"I usually set personal goals and feel frustrated when I don't meet them. The AI asked, 'Why do you set such standards?' which made me select it because it questioned something I had always taken for granted, like the fact that the Earth is round. It was an idea I had never thought about, and that's why it stood out to me."* In contrast, some participants avoided more abstract or challenging questions, opting for those that *"felt easier to answer based on concrete experiences (P5)."* To balance the cognitive load, several participants voluntarily alternated between engaging with challenging, thought-provoking questions and those they found more manageable.

*6.3.3 Objectification and Reiteration through AI Summary.* Participants generally viewed the AI summary to bring clarity and objectivity to their often complex and unorganized thoughts by distilling the fragmented responses into an organized narrative. In turn, the AI summary seemed to foster iterative reflection; twelve out of 19 participants (63%) revisited the Exploration page, explored more, and went back to the AI summary page to check the new summary reflecting additional responses (See dark blue bars in Figure 6). P12 remarked, *"I liked the summary, but I was curious how it might turn out if I revisited it, so I kept checking. After seeing the summary again, I thought of things I had missed and wrote them down, then went back to see how the updated summary looked. This back-and-forth really helped me organize my thoughts."* This iterative process, where participants reviewed and refined their reflections, enhanced their engagement and allowed for further self-exploration.

Some participants also reported that the AI summary helped them face uncomfortable or avoided thoughts. P18 recalled that she initially wrote, *"Watching YouTube videos makes me more tired,"* and seeing the AI repeat it in the summary forced her to confront this feeling more directly. The AI's neutral tone in the summary helped participants face these challenging thoughts without feeling overly judged, as P2 also described, *"It's something I already know, but the AI organizes my situation objectively. If a person did this, it might feel uncomfortable, but with AI, it doesn't bother me."*

## 6.4 Supporting Self-Expression

On the question panels, ExploreSelf provided adaptive guidance, including keywords and comments. In this section, we cover how these elements of guidance helped participants' exploration and reflection process.

*6.4.1 Enriching and Rethinking Expressions through Keywords.* Participants could toggle on the keywords to reveal them, and seventeen out of 19 (89%) participants saw keywords more than once during the exploration phase (see 'Total # of keywords' in Table 4), but with varying levels of engagement with them. Some participants received keywords as a valuable tool for articulating their thoughts, while others either used it selectively or chose not to rely on them. Many participants used the keywords as prompts to help express their emotions and ideas more clearly. For instance, P7 mentioned while reflecting on emotions after a conflict with

their children, he initially thought of simplistic terms like *"anger"* or *"sadness."* However, the keywords suggested deeper and more nuanced emotions like *"regret"* and *"powerlessness,"* which broadened their emotional vocabulary and helped express himself more fully. Similarly, P14, emphasizing that her age is over 60, mentioned that *"As I age, I struggle to think of the right words and expressions, but seeing the keywords made it easier to express myself."*

Some participants used the keywords as a way to explore new perspectives or deepen their reflections. P8, for instance, frequently clicked *See More Keywords* by mentioning, *"the AI was smarter, showing me words I couldn't have thought of."* The keywords would prompt participants to think beyond their initial responses and anchor their reflections. P13 also mentioned that keywords sparked associations with aspects of the matter that she had not thought of in the first place. At the same time, some participants reported that the keywords mitigated the mental burden while organizing negative thoughts for answering. P4, who had experienced trauma, found the keywords helpful in guiding him through some of his painful memories. He noted that the keywords made it less emotionally overwhelming to revisit these thoughts because EXPLORESELF offered possible expressions they could choose from, making the process feel more manageable.

*6.4.2 Iteratively Developing and Refining Answers through Comments.* Participants were exposed to one comment per question by default, and fifteen out of 19 (79%) requested additional comments more than once during the exploration phase (see 'Total # of comment requests' in Table 4), with varying number of requests per participant. Initial impressions of the comments seemed to shape how participants continued to use or avoid the feature. For instance, P17 mentioned that while they initially had low expectations, receiving an encouraging comment led them to revisit the feature more often, curious about what other positive feedback they might receive. On the other hand, P9, who has requested comment only once, noted that after receiving a comment that felt irrelevant to their context, they decided not to use the feature again.

Many participants appreciated the responsiveness of the comments, which provided a sense of dynamic interaction. For instance, P19 noted that the comments felt like having a *"running mate,"* offering feedback that helped him stay engaged. P13 similarly expressed curiosity about how the comments would respond, motivating her to continue interacting. P11 described how the comments helped refine his thoughts, iteratively guiding him from a single word, *"financial stability"*, to a more developed sentence. The comment prompted him to *"think about what you ultimately want from financial stability,"* which led to a detailed response about *"owning property, a car, and a happy family."*

Comments generally provided participants with serendipitous insights through a more detailed message, and participants thought a comment was particularly insightful when it acknowledged their feelings or thoughts with clear rationales: *"[...] The fact that it acknowledged my feelings as natural was really meaningful to me (P7)."* Although the encouraging tone of comments often engaged participants in the exploration, not all experiences were positive. For example, P16 felt that some comments were overly upbeat and didn't match the gravity of her situation, causing a sense of emotional disconnection.

## 7 Discussion

In this section, we discuss lessons learned from the design and implementation of EXPLORESELF as well as from the user study. We also reflect on the implications for better supporting LLM-guided self-reflection in a long-term and inclusive setting.

### 7.1 Supporting User-driven Exploration for Mental Well-being

EXPLORESELF facilitated user control over their reflective journey by offering multiple options and layers of adaptive support that encouraged users to navigate their thoughts and emotions. This scaffolded structure allowed participants to tailor their exploration pathways, balancing their cognitive and emotional burdens while facing various aspects of their personal challenges. For example, participants chose to engage more deeply with a theme through Socratic questions or pivoted to new themes when they felt ready, demonstrating varied patterns of exploration (see Figure 7). Consequently, participants showed a significant increase in perceived agency after engaging with EXPLORESELF, underscoring its potential to facilitate for users to manage their reflective journey.

Our main design goal with EXPLORESELF was to prioritize user agency without directing content, while providing guidance. This goal resulted in the emphasis on *exploration*, leading to design choices that are different from prior LLM-assisted reflection tools, such as DiaryMate [48] and MindfulDiary [47]. For example, DiaryMate employed the concept of writing assistant [55] and used LLMs to generate diary sentences given user inputs like keywords and the model temperature [48]. While this approach helped participants complete well-written journal entries, the quality and content of LLM generations would largely influence their reflection and even sometimes interrupt their own stream of thought [48]. MindfulDiary incorporated a conversational agent, loosely mimicking a counseling interview with users [47]. Although the conversational approach effectively demonstrated care and empathy, the LLM-driven agent mostly held the initiative of conversation and the turn-taking interaction focused the user's attention on a single question presented at each turn [47]. EXPLORESELF differs from previous work in that, unlike DiaryMate, it offers users choices about what to explore (*i.e.*, themes and questions) rather than generating narratives on their behalf. Additionally, users could determine their level of engagement with each topic through an interface designed to minimize single-path conversational affordances, as in MindfulDiary. This approach encouraged users to 'contemplate' their exploration, enabling user-directed control over the process.

In this work, we explored a novel design approach to invite user agency throughout a self-paced exploration for mental well-being. EXPLORESELF targeted lay individuals who may not require clinical diagnoses but any of those who would willingly seek assistance in navigating personal challenges. In this context, its approach follows the principles of psychotherapy, particularly its emphasis on client autonomy [102]. Guided by self-determination theory [24], it is widely agreed that supporting clients in autonomously exploring and sustaining a process of change is crucial for a successful outcome [85, 100, 102]. Similarly, ExploreSelf provided an environment where users can process their challenges at their own pace and on their own terms, which mirrors the role of therapist—fostering a

space where individuals feel empowered to explore their emotions, identify goals, and develop strategies for personal growth [101]—illustrating the potential of designing adaptive guidance with LLMs.

## 7.2 Design Considerations for Long-term, Multi-session Interaction

At an early stage of investigating the feasibility of leveraging LLM-guided exploration of personal challenges, our work was based on a single-session lab study. That being said, in the debriefing, all participants except one expressed willingness to engage with ExploreSelf continuously in the future, particularly in circumstances where they would need to process complex thoughts or emotions. For example, P14 imagined the system could help users recognize recurring themes, providing deeper insights into long-standing thought patterns and behaviors. Their feedback underscores ExploreSelf's potential to serve as a channel for people to engage with regularly to process their personal challenges.

Supporting long-term engagement with ExploreSelf—revisiting ExploreSelf in multiple sessions over time—requires a new set of design considerations. The current version of ExploreSelf expands the branches of reflection rooted in the initial narrative, which the user provides in a prelude to the exploration. Intuitively treating this unit as a session, users may create multiple sessions or append more themes to a previously-explored session. Notably, the sessions of the same user may be interconnected, sharing common aspects of and sometimes influencing one another. For example, a series of personal challenges can lead to the discovery of an overarching theme. Also to ensure accuracy and reliability, the system should incorporate sort of *long-term memory* (*c.f.*, [5, 42]), where it leverages the information derived from past sessions for the current session. For example, the system may memorize the interpersonal relationships of the user and consult it when generating themes and questions accordingly. However, simply stitching the entire interaction history for model input is neither scalable nor reliable [5], and could lead to inaccurate or irrelevant summaries and guidance. This calls for future research in designing how the system constructs and synthesizes the user knowledge that is noteworthy.

Also, prolonged user engagement with the system will also increase the volume of interaction history to be presented. Therefore, interaction techniques such as semantic zooming [10] and the multi-level summarization approach would be necessary. The long-term data will also pose opportunities to present other supplementary data, such as recurring questions or themes that the user keeps avoiding exploring: P13 imagined, "*I noticed that I have avoided the questions on a specific topic. Seeing this allowed me to better understand myself about what I don't want to face.*" Supporting such 'byproduct' reflections calls for further investigation from the perspectives of the personal informatics systems [15, 50, 61].

To comprehensively support people in addressing personal challenges, future versions of ExploreSelf should go beyond the exploration phase. Our user study results indicate that about 30 minutes of exploration fostered a certain level of sensemaking about the self and the subject matter. According to the transtheoretical model [93], the exploration mostly supports individuals to reach the *contemplation* stage, where they recognize what to change to address their challenge. However, different types of guidance and interaction may

be more effective for those who have progressed to the *preparation* stage, where they began taking small steps towards change.

## 7.3 Role and Stance of AI in Guiding Reflection

Although we minimized conversational affordances from the interface, ExploreSelf still conveyed messages in human language with a tone through comments. Within the current scope of investigation, we followed a broadly applicable guideline and instructed the model to generate empathetic and supportive comments. However, the positive tone of comments was sometimes perceived as inconsiderate of the gravity of the participants' concerns. This underscores the need for a more nuanced design of the AI's role and stance embedded in its messages. While ExploreSelf is not intended to mimic psychotherapy, the therapists from our formative interviews highlighted that balancing both empathy and confrontation is often a helpful therapeutic tactic, pointing out that they do not always show agreement and support to their clients in leading them to positive change. Notably, counselors sometimes *confront* their clients with discrepancies in their perception of reality [2], which may lead to resistance [96].

Applied to ExploreSelf, the AI may take a stance that properly facilitates reflection on the fly that is aligned with these principles. To enable users to engage with confronting guidance when they are mentally ready, the system may provide a set of different stances that users can select as needed. However, LLMs, particularly those fine-tuned with human alignment techniques like RLHF [82], typically exhibit supportive, empathetic, and cooperative behaviors [57, 119, 124] and this tendency is difficult to override by prompting alone [62, 123]. To demonstrate a confronting stance, the system may leverage customized LLMs fine-tuned with annotated transcripts from psychotherapy sessions such as Motivational Interviewing [96]. Here, users' interaction logs would also play a key role in assessing their resistance to specific themes. For instance, P12 and P14 noted that reviewing unselected questions might uncover the aspects of the challenge that he often avoids contemplating. A future version of ExploreSelf can analyze the patterns in these *roads not taken* and determine the level of confrontational guidance provided to the user.

## 7.4 Supporting Severe and Sensitive Challenges

Observing participants exploring diverse personal challenges of varying severity, we identified implications for supporting individuals with severe challenges, including those related to mental illness, such as depression, traumas, or other emotional distress. For individuals experiencing severe apathy, commonly seen in conditions like depression, self-directed exploration may be overwhelming [47, 109]. This echoes the feedback from experts in our formative interviews, who noted that individuals with severe emotional numbness may struggle to engage meaningfully with open-ended reflective tools. In such cases, more proactive guidance can be more effective than a high degree of choice.

Similarly, reflecting on one's trauma necessitates careful guidance due to its emotional intensity involved [71, 91]. To keep the exploration within the emotionally safe boundary, ExploreSelf may receive traumatic keywords or precautions from users in advance to consider them when handling the emotionally charged

material. Although EXPLORESELF allows users to decide when to stop drilling into certain themes, individuals exploring a traumatic challenge may become overly involved in the exploration. Future work can investigate whether and how the system should intervene during exploration. However, simply terminating a session, particularly when users lack alternative resources, raises a lot of ethical concerns about potentially leaving individuals without support in critical moments. Instead, it is essential to develop guardrails that facilitate navigation for vulnerable users while ensuring their safety. This is especially important given that many individuals increasingly turn to accessible LLM interfaces for coping with mental health challenges in the absence of other available options [113].

Cultural awareness is crucial for addressing deeply complex and sensitive challenges, as cultural consideration is an integral aspect of mental health support [7, 17, 58, 87]. Considering the technical reliability and availability, we selected OpenAI GPT—an LLM trained with datasets dominated by English and Western cultural references—as the underlying model for our generative pipelines in Korean. Although we did not receive participants' feedback that the AI generations lacked sensitivity to Korean cultural norms, the system may still reflect certain Western cultural norms and biases [86, 114] when handling challenges specific to non-Western contexts. This limitation can be particularly problematic when engaging with marginalized communities, for whom issues such as racism or sexism are not mere distortions but ongoing harms [35, 67]. Employing LLMs aligned with the language and cultural context of the target population (*e.g.*, HyperCLOVA [46] in Korean) could better reflect culturally relevant references [75, 105]. However, the disparity in the model performance of LLMs across different languages and regions remains an open challenge.

## 7.5 Limitations and Future Work

In this section, we discuss the limitations of our study that may impact the generalizability of the findings. First, our study participants might not have fully disclosed their innermost challenges, given it is a research study setting. We tried to address this issue by: (1) minimizing the invasiveness of observation by creating breakout rooms where participants can have their privacy in exploration; and (2) collecting and observing only the interaction logs, not the actual problem narratives to respect their privacy. To further ensure their safety, we limited the time of exploration so that the study would not get too emotionally charged for our participants, and prepared safety protocols in case of emergency. We also acknowledge that collecting data from more ecologically valid settings (*e.g.*, field deployment) may yield different distributions of topics and exploration patterns. Furthermore, although we aimed to recruit participants from diverse backgrounds, all participants are residents in South Korea. Populations with different socio-cultural contexts and AI literacy may engage in different types of personal challenges using EXPLORESELF.

## 8  Conclusion

We presented EXPLORESELF, an LLM-driven application that supports processing personal challenges by exploring related themes and questions. EXPLORESELF enables people to take the initiative of exploration, thereby maintaining their agency and proactively selecting what aspects they want to explore. To examine how people use EXPLORESELF to explore and reflect on their own challenges, we conducted an exploratory user study with 19 individuals. Participants successfully adopted the concept of proactive exploration interactions that EXPLORESELF offered, showing various patterns. Our study revealed that providing multiple options with adaptive guidance enhanced participants' perceived agency, and the large language model played a key role by providing wordings that acknowledged the participants' language in detail. We highlighted several areas for future research, including designing LLM-guided self-reflection systems that balance autonomy and intervention, technical and design considerations for long-term reflective practices, and incorporating cultural and emotional sensitivity to address diverse user needs and challenges. In summary, our work contributes to the growing body of LLM-driven mental well-being support, demonstrating the feasibility and usefulness of flexible user interface inspired by reflective writing.

## References

[1] Kathleen Adams. 2009. *Journal to the self: Twenty-two paths to personal growth-open the door to self-understanding bu writing, reading, and creating a journal of your life.* Hachette UK.

[2] Susan C Anderson. 1968. Effects of Confrontation by High- and Low-functioning Therapists. *Journal of Counseling Psychology* 15, 5p1 (1968), 411.

[3] Matthew Anders Andersson and Colleen S Conley. 2008. Expecting to heal through self-expression: A perceived control theory of writing and health. *Health Psychology Review* 2, 2 (2008), 138–162.

[4] Anthropic. 2024. The Claude 3 Model Family: Opus, Sonnet, Haiku. Retrieved September 11, 2024 from https://www-cdn.anthropic.com/de8ba9b01c9ab7cbabf5c33b80b7bbc618857627/Model_Card_Claude_3.pdf

[5] Sanghwan Bae, Donghyun Kwak, Soyoung Kang, Min Young Lee, Sungdong Kim, Yuin Jeong, Hyeri Kim, Sang-Woo Lee, Woomyoung Park, and Nako Sung. 2022. Keep Me Updated! Memory Management in Long-term Conversations. In *Findings of the Association for Computational Linguistics: EMNLP 2022*, Yoav Goldberg, Zornitsa Kozareva, and Yue Zhang (Eds.). Association for Computational Linguistics, Abu Dhabi, United Arab Emirates, 3769–3787. https://doi.org/10.18653/v1/2022.findings-emnlp.276

[6] Karen A Baikie and Kay Wilhelm. 2005. Emotional and physical health benefits of expressive writing. *Advances in psychiatric treatment* 11, 5 (2005), 338–346.

[7] Luke Balcombe and Diego De Leo. 2022. Human-computer interaction in digital mental health. In *Informatics*, Vol. 9. MDPI, 14.

[8] Albert Bandura. 1991. Social cognitive theory of self-regulation. *Organizational behavior and human decision processes* 50, 2 (1991), 248–287.

[9] Oliver Blunk and Michael Prilla. 2017. Prompting to support reflection: A workplace study. In *Data Driven Approaches in Digital Education: 12th European Conference on Technology Enhanced Learning, EC-TEL 2017, Tallinn, Estonia, September 12–15, 2017, Proceedings 12*. Springer, 367–372.

[10] Maged N Kamel Boulos. 2003. The use of interactive graphical maps for browsing medical/health Internet information resources. *International Journal of Health Geographics* 2, 1 (2003), 1. https://doi.org/10.1186/1476-072x-2-1

[11] Virginia Braun and Victoria Clarke. 2006. Using Thematic Analysis in Psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. https://doi.org/10.1191/1478088706qp063oa

[12] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel Ziegler, Jeffrey Wu, Clemens Winter, Chris

Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. Language Models are Few-Shot Learners. In *Advances in Neural Information Processing Systems (NeurIPS '20)*, H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin (Eds.), Vol. 33. Curran Associates, Inc., 1877–1901. https://proceedings.neurips.cc/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64a-Paper.pdf

[13] Alexandra G Buchlmayer, Lucas E Snyder, Ian Sotomayor, Arielle C Smith, Laura K Jans, Jessica L Schleider, Myriam El Khoury-Malhame, and Cosette Fox. 2024. Effects of an Online Growth Mindset Single-Session Intervention on Hopelessness, Self-Hate, and Perception of Control in Arabic Adolescents. *Psi Chi Journal of Psychological Research* 29, 1 (2024).

[14] Yu Ying Chiu, Ashish Sharma, Inna Wanyin Lin, and Tim Althoff. 2024. A Computational Framework for Behavioral Assessment of LLM Therapists. arXiv:2401.00820 [cs.CL] https://arxiv.org/abs/2401.00820

[15] Eun Kyoung Choe, Bongshin Lee, Haining Zhu, Nathalie Henry Riche, and Dominikus Baur. 2017. Understanding Self-Reflection: How People Reflect on Personal Data Through Visual Data Exploration. *Proc. EAI International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth '17)* (2017).

[16] Yu-Hee Choi, Hee-Kyung Lee, and DG Lee. 2008. Validation of the Korean version of Snyder's dispositional hope scale. *Korean Journal of Social and Personality Psychology* 22, 2 (2008), 1–16.

[17] John Chambers Christopher. 1999. Situating psychological well-being: Exploring the cultural roots of its theory and research. *Journal of Counseling & Development* 77, 2 (1999), 141–152.

[18] Gavin I Clark and Sarah J Egan. 2015. The Socratic method in cognitive behavioural therapy: a narrative review. *Cognitive Therapy and Research* 39 (2015), 863–879.

[19] Timothy J Cleary and Barry J Zimmerman. 2004. Self-regulation empowerment program: A school-based program to enhance self-regulated and self-motivated cycles of student learning. *Psychology in the Schools* 41, 5 (2004), 537–550.

[20] Sharon Danoff-Burg, Catherine E Mosher, Asani H Seawell, and John D Agee. 2010. Does narrative writing instruction enhance the benefits of expressive writing? *Anxiety, Stress, & Coping* 23, 3 (2010), 341–352.

[21] Elizabeth A Davis. 2003. Prompting middle school science students for productive reflection: Generic and directed prompts. *The Journal of the Learning Sciences* 12, 1 (2003), 91–142.

[22] Edward L Deci and Richard M Ryan. 1987. The support of autonomy and the control of behavior. *Journal of personality and social psychology* 53, 6 (1987), 1024.

[23] 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.

[24] Edward L Deci and Richard M Ryan. 2013. *Intrinsic motivation and self-determination in human behavior*. Springer Science & Business Media.

[25] Benjamin Dominguez, Pablo Valderrama, María de los Angeles Meza, Sara Lidia Perez, Amparo Silva, Gloria Martinez, Victor Manuel Mendez, and Yolanda Olvera. 1995. The roles of disclosure and emotional reversal in clinical practice. (1995).

[26] Miriam Evans and Alyssa R Boucher. 2015. Optimizing the power of choice: Supporting student autonomy to foster motivation and engagement in learning. *Mind, Brain, and Education* 9, 2 (2015), 87–91.

[27] Nancy J Evans. 2011. Authoring Your Life: Developing an Internal Voice to Navigate Life's Challenges. *Journal of College Student Development* 52, 2 (2011), 248–250.

[28] Anya S Evmenova, Kelley Regan, Reagan Mergen, and Roba Hrisseh. 2024. Improving Writing Feedback for Struggling Writers: Generative AI to the Rescue? *TechTrends* (2024), 1–13.

[29] Erin M Fekete and Nathan T Deichert. 2022. A brief gratitude writing intervention decreased stress and negative affect during the COVID-19 pandemic. *Journal of Happiness Studies* 23, 6 (2022), 2427–2448.

[30] OpenJS Foundation. 2024. Express–Node.js Web Application Framework. Retrieved Sep 10, 2024 from https://expressjs.com/

[31] Patricia Frazier, Nora Keenan, Samantha Anders, Sulani Perera, Sandra Shallcross, and Samuel Hintz. 2011. Perceived past, present, and future control and adjustment to stressful life events. *Journal of personality and social psychology* 100, 4 (2011), 749.

[32] Google, Inc. 2023. Bard - Chat Based AI Tool from Google, Powered by PaLM 2. Retrieved Sep 10, 2024 from https://bard.google.com/

[33] Juhye Ha, Hyeon Jeon, Daeun Han, Jinwook Seo, and Changhoon Oh. 2024. CloChat: Understanding How People Customize, Interact, and Experience Personas in Large Language Models. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 305, 24 pages. https://doi.org/10.1145/3613904.3642472

[34] John H Harvey, Terri L Orbuch, and Kelly Fink. 1990. The social psychology of account-making: Meaning, hope, and generativity. *New Zealand Journal of Psychology* 19, 2 (1990), 46–57.

[35] Sally Haslanger. 2017. Racism, ideology, and social movements. (2017).

[36] Steven C Hayes and Heather Pierson. 2005. *Acceptance and Commitment Therapy*. Springer.

[37] Steve Hodges, Lyndsay Williams, Emma Berry, Shahram Izadi, James Srinivasan, Alex Butler, Gavin Smyth, Narinder Kapur, and Ken Wood. 2006. SenseCam: A retrospective memory aid. In *UbiComp 2006: Ubiquitous Computing: 8th International Conference, UbiComp 2006 Orange County, CA, USA, September 17-21, 2006 Proceedings 8*. Springer, 177–193.

[38] Katherine Hoogesteyn, Ewout H Meijer, and Aldert Vrij. 2020. Detainee and layperson's expectations and preferences regarding police interview rooms. *Plos one* 15, 11 (2020), e0241683.

[39] Delaura L Hubbs and Charles F Brand. 2005. The paper mirror: Understanding reflective journaling. *Journal of Experiential Education* 28, 1 (2005), 60–71.

[40] Inflection AI. 2023. Pi, your personal AI. Retrieved Sep 10, 2024 from https://pi.ai/talk

[41] Sooyeon Jeong and Cynthia Lynn Breazeal. 2016. Improving smartphone users' affect and wellbeing with personalized positive psychology interventions. In *Proceedings of the fourth international conference on human agent interaction*. 131–137.

[42] Eunkyung Jo, Yuin Jeong, Sohyun Park, Daniel A. Epstein, and Young-Ho Kim. 2024. Understanding the Impact of Long-Term Memory on Self-Disclosure with Large Language Model-Driven Chatbots for Public Health Intervention. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 440, 21 pages. https://doi.org/10.1145/3613904.3642420

[43] Pelin Karaturhan, Ecem Arıkan, Pelin Durak, Asim Evren Yantac, and Kemal Kuscu. 2022. Combining momentary and retrospective self-reflection in a mobile photo-based journaling application. In *Nordic Human-Computer Interaction Conference*. 1–12.

[44] Todd B Kashdan and Jonathan Rottenberg. 2010. Psychological flexibility as a fundamental aspect of health. *Clinical psychology review* 30, 7 (2010), 865–878.

[45] Douglas F Kauffman, Xun Ge, Kui Xie, and Ching-Huei Chen. 2008. Prompting in web-based environments: Supporting self-monitoring and problem solving skills in college students. *Journal of Educational Computing Research* 38, 2 (2008), 115–137.

[46] Boseop Kim, HyoungSeok Kim, Sang-Woo Lee, Gichang Lee, Donghyun Kwak, Jeon Dong Hyeon, Sunghyun Park, Sungju Kim, Seonhoon Kim, Dongpil Seo, Heungsub Lee, Minyoung Jeong, Sungjae Lee, Minsub Kim, Suk Hyun Ko, Seokhun Kim, Taeyong Park, Jinuk Kim, Soyoung Kang, Na-Hyeon Ryu, Kang Min Yoo, Minsuk Chang, Soobin Suh, Sookyo In, Jinseong Park, Kyungduk Kim, Hiun Kim, Jisu Jeong, Yong Goo Yeo, Donghoon Ham, Dongju Park, Min Young Lee, Jaewook Kang, Inho Kang, Jung-Woo Ha, Woomyoung Park, and Nako Sung. 2021. What Changes Can Large-scale Language Models Bring? Intensive Study on HyperCLOVA: Billions-scale Korean Generative Pretrained Transformers. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Online and Punta Cana, Dominican Republic, 3405–3424. https://doi.org/10.18653/v1/2021.emnlp-main.274

[47] 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 CHI Conference on Human Factors in Computing Systems*. 1–20.

[48] 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 CHI Conference on Human Factors in Computing Systems*. 1–15.

[49] Youngsuk Kim. 2005. *Effects of expressive writing among Mexican and Korean bilinguals on social, physical, and mental well-being*. The University of Texas at Austin.

[50] Young-Ho Kim, Bongshin Lee, Arjun Srinivasan, and Eun Kyoung Choe. 2021. Data@Hand: Fostering Visual Exploration of Personal Data on Smartphones Leveraging Speech and Touch Interaction. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. ACM, New York, NY, USA, Article 462, 17 pages. https://doi.org/10.1145/3411764.3445421

[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] Harsh Kumar, Ruiwei Xiao, Benjamin Lawson, Ilya Musabirov, Jiakai Shi, Xinyuan Wang, Huayin Luo, Joseph Jay Williams, Anna N. Rafferty, John Stamper, and Michael Liut. 2024. Supporting Self-Reflection at Scale with Large Language Models: Insights from Randomized Field Experiments in Classrooms. In *Proceedings of the Eleventh ACM Conference on Learning @ Scale* (Atlanta, GA, USA) *(L@S '24)*. Association for Computing Machinery, New York, NY, USA, 86–97. https://doi.org/10.1145/3657604.3662042

[53] Inc. LangChain. 2024. LangChain: Applications that Can Reason. Retrieved Sep 10, 2024 from https://www.langchain.com/

[54] Thomas A Langens and Julia Schüler. 2007. Effects of written emotional expression: The role of positive expectancies. *Health Psychology* 26, 2 (2007), 174.

[55] Mina Lee, Katy Ilonka Gero, John Joon Young Chung, Simon Buckingham Shum, Vipul Raheja, Hua Shen, Subhashini Venugopalan, Thiemo Wambsganss, David Zhou, Emad A. Alghamdi, Tal August, Avinash Bhat, Madiha Zahrah Choksi, Senjuti Dutta, Jin L.C. Guo, Md Naimul Hoque, Yewon Kim, Simon Knight, Seyed Parsa Neshaei, Antonette Shibani, Disha Shrivastava, Lila Shroff, Agnia Sergeyuk, Jessi Stark, Sarah Sterman, Sitong Wang, Antoine Bosselut, Daniel Buschek, Joseph Chee Chang, Sherol Chen, Max Kreminski, Joonsuk Park, Roy Pea, Eugenia Ha Rim Rho, Zejiang Shen, and Pao Siangliulue. 2024. A Design Space for Intelligent and Interactive Writing Assistants. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 1054, 35 pages. https://doi.org/10.1145/3613904.3642697

[56] Mina Lee, Percy Liang, and Qian Yang. 2022. CoAuthor: Designing a Human-AI Collaborative Writing Dataset for Exploring Language Model Capabilities. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 388, 19 pages. https://doi.org/10.1145/3491102.3502030

[57] Yoon Kyung Lee, Jina Suh, Hongli Zhan, Junyi Jessy Li, and Desmond C. Ong. 2024. Large Language Models Produce Responses Perceived to be Empathic. arXiv:2403.18148 [cs.CL] https://arxiv.org/abs/2403.18148

[58] Robert W Lent. 2004. Toward a unifying theoretical and practical perspective on well-being and psychosocial adjustment. *Journal of counseling psychology* 51, 4 (2004), 482.

[59] SJ Lepore. 2002. The writing cure: How expressive writing promotes health and emotional well-being. *American Psychological Association* (2002).

[60] Stephen J Lepore and Wendy Kliewer. 2020. Expressive writing and health. In *Encyclopedia of behavioral medicine*. Springer, 817–823.

[61] Ian Li, Anind K. Dey, and Jodi Forlizzi. 2011. Understanding My Data, Myself: Supporting Self-reflection with Ubicomp Technologies. In *Proceedings of the 13th International Conference on Ubiquitous Computing* (Beijing, China) *(UbiComp '11)*. ACM, New York, NY, USA, 405–414. https://doi.org/10.1145/2030112.2030166

[62] Wenkai Li, Jiarui Liu, Andy Liu, Xuhui Zhou, Mona Diab, and Maarten Sap. 2024. BIG5-CHAT: Shaping LLM Personalities Through Training on Human-Grounded Data. arXiv:2410.16491 [cs.CL] https://arxiv.org/abs/2410.16491

[63] Qian Lu, Dianhan Zheng, Lucy Young, Marjorie Kagawa-Singer, and Alice Loh. 2012. A pilot study of expressive writing intervention among Chinese-speaking breast cancer survivors. *Health Psychology* 31, 5 (2012), 548.

[64] Haiwei Ma, Sunny Parawala, and Svetlana Yarosh. 2021. Detecting Expressive Writing in Online Health Communities by Modeling Aggregated Empirical Data. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–32.

[65] Zilin Ma, Yiyang Mei, and Zhaoyuan Su. 2023. Understanding the benefits and challenges of using large language model-based conversational agents for mental well-being support. *AMIA Annu. Symp. Proc.* 2023 (2023), 1105–1114.

[66] David Manier and Adelaida Olivares. 2005. Who Benefits from Expressive Writing? Moderator Variables Affecting Outcomes of Emotional Disclosure Interventions. *Counseling & Clinical Psychology Journal* 2, 1 (2005).

[67] Mary Kathryn McGowan. 2019. *Just words: on speech and hidden harm*. Oxford University Press.

[68] Liang Meng and Qingguo Ma. 2015. Live as we choose: the role of autonomy support in facilitating intrinsic motivation. *International Journal of Psychophysiology* 98, 3 (2015), 441–447.

[69] Meta. 2024. React: The library for web and native user interfaces. Retrieved Sep 10, 2024 from https://react.dev

[70] Microsoft. 2024. TypeScript. Retrieved Sep 10, 2024 from https://www.typescriptlang.org

[71] Madelyn Miller. 2001. Creating a safe frame for learning: Teaching about trauma and trauma treatment. *Journal of Teaching in Social Work* 21, 3-4 (2001), 159–176.

[72] William Miller. 2014. Interactive journaling as a clinical tool. *Journal of mental health counseling* 36, 1 (2014), 31–42.

[73] MongoDB, Inc. 2024. MongoDB: The Developer Data Platform | MongoDB. Retrieved Sep 10, 2024 from https://www.mongodb.com/

[74] Jase Moussa-Inaty. 2015. Reflective writing through the use of guiding questions. *International Journal of Teaching and Learning in Higher Education* 27, 1 (2015), 104–113.

[75] Tarek Naous, Michael Ryan, Alan Ritter, and Wei Xu. 2024. Having Beer after Prayer? Measuring Cultural Bias in Large Language Models. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Lun-Wei Ku, Andre Martins, and Vivek Srikumar (Eds.). Association for Computational Linguistics, Bangkok, Thailand, 16366–16393. https://aclanthology.org/2024.acl-long.862

[76] Carla Nave, Nuno Correia, and Teresa Romão. 2016. Exploring emotions through painting, photography and expressive writing: An early experimental user study. In *Proceedings of the 13th International Conference on Advances in Computer Entertainment Technology*. 1–8.

[77] Subigya Nepal, Arvind Pillai, William Campbell, Talie Massachi, Eunsol Soul Choi, Xuhai Xu, Joanna Kuc, Jeremy F Huckins, Jason Holden, Colin Depp, Nicholas Jacobson, Mary P Czerwinski, Eric Granholm, and Andrew Campbell. 2024. Contextual AI Journaling: Integrating LLM and Time Series Behavioral Sensing Technology to Promote Self-Reflection and Well-being using the MindScape App. In *Extended Abstracts of the 2024 CHI Conference on Human Factors in Computing Systems (CHI EA '24)*. Association for Computing Machinery, New York, NY, USA, Article 86, 8 pages. https://doi.org/10.1145/3613905.3650767

[78] Kathleen O'Leary, Stephen M Schueller, Jacob O Wobbrock, and Wanda Pratt. 2018. "Suddenly, we got to become therapists for each other" Designing Peer Support Chats for Mental Health. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–14.

[79] OpenAI. 2024. ChatGPT: Optimizing Language Models for Dialogue. Retrieved Sep 10, 2024 from https://openai.com/blog/chatgpt/

[80] OpenAI. 2024. OpenAI API. Retrieved Sep 10, 2024 from https://openai.com/api/

[81] OpenAI, Inc. 2023. Replika. https://replika.com/

[82] Long Ouyang, Jeff Wu, Xu Jiang, Diogo Almeida, Carroll L. Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul Christiano, Jan Leike, and Ryan Lowe. 2022. Training language models to follow instructions with human feedback. arXiv:2203.02155 [cs.CL] https://arxiv.org/abs/2203.02155

[83] SoHyun Park, Anja Thieme, Jeongyun Han, Sungwoo Lee, Wonjong Rhee, and Bongwon Suh. 2021. "I wrote as if I were telling a story to someone I knew.": Designing Chatbot Interactions for Expressive Writing in Mental Health. In *Proceedings of the 2021 ACM Designing Interactive Systems Conference*. 926–941.

[84] Samir Passi and Mihaela Vorvoreanu. 2022. Overreliance on AI Literature Review. *Microsoft Research* (2022).

[85] Erika A Patall, Harris Cooper, and Jorgianne Civey Robinson. 2008. The effects of choice on intrinsic motivation and related outcomes: a meta-analysis of research findings. *Psychological bulletin* 134, 2 (2008), 270.

[86] Siddhesh Pawar, Junyeong Park, Jiho Jin, Arnav Arora, Junho Myung, Srishti Yadav, Faiz Ghifari Haznitrama, Inhwa Song, Alice Oh, and Isabelle Augenstein. 2024. Survey of Cultural Awareness in Language Models: Text and Beyond. *arXiv preprint arXiv:2411.00860* (2024).

[87] Sachin R Pendse, Naveena Karusala, Divya Siddarth, Pattie Gonsalves, Seema Mehrotra, John A Naslund, Mamta Sood, Neha Kumar, and Amit Sharma. 2019. Mental health in the global south: challenges and opportunities in HCI for development. In *Proceedings of the 2nd ACM SIGCAS Conference on Computing and Sustainable Societies*. 22–36.

[88] James W Pennebaker. 1993. Putting stress into words: Health, linguistic, and therapeutic implications. *Behaviour research and therapy* 31, 6 (1993), 539–548.

[89] James W Pennebaker. 1997. Writing about emotional experiences as a therapeutic process. *Psychological science* 8, 3 (1997), 162–166.

[90] James W Pennebaker, Janice K Kiecolt-Glaser, and Ronald Glaser. 1988. Disclosure of traumas and immune function: health implications for psychotherapy. *Journal of consulting and clinical psychology* 56, 2 (1988), 239.

[91] Stuart D Perlman. 1998. *The therapist's emotional survival: Dealing with the pain of exploring trauma*. Jason Aronson, Incorporated.

[92] Aleksandar Petrov, Emanuele La Malfa, Philip H. S. Torr, and Adel Bibi. 2023. Language Model Tokenizers Introduce Unfairness Between Languages. In *Advances in Neural Information Processing Systems (NeurIPS '23)*. https://arxiv.org/abs/2305.15425

[93] James O. Prochaska and Wayne F. Velicer. 1997. The Transtheoretical Model of Health Behavior Change. *American Journal of Health Promotion* 12, 1 (1997), 38–48. https://doi.org/10.4278/0890-1171-12.1.38 arXiv:https://doi.org/10.4278/0890-1171-12.1.38 PMID: 10170434.

[94] Ira Progoff. 1977. *At a journal workshop: The basic text and guide for using the Intensive Journal*. Dialogue House Library.

[95] Kelley S Regan, Margo A Mastropieri, and Thomas E Scruggs. 2005. Promoting expressive writing among students with emotional and behavioral: Disturbance via dialogue journals. *Behavioral disorders* 31, 1 (2005), 33–50.

[96] Ken Resnicow and Fiona McMaster. 2012. Motivational Interviewing: Moving from Why to How with Autonomy Support. *International Journal of Behavioral Nutrition and Physical Activity* 9 (2012), 1–9.

[97] Josephine P Riddle, HE Smith, and CJ Jones. 2016. Does written emotional disclosure improve the psychological and physical health of caregivers? A systematic review and meta-analysis. *Behaviour research and therapy* 80 (2016), 23–32.

[98] Chiara Ruini and Cristina C Mortara. 2022. Writing technique across psychotherapies—from traditional expressive writing to new positive psychology interventions: A narrative review. *Journal of Contemporary Psychotherapy* (2022), 1–12.

[99] Richard M Ryan and Edward L Deci. 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist* 55, 1 (2000), 68.

[100] Richard M Ryan and Edward L Deci. 2008. A self-determination theory approach to psychotherapy: The motivational basis for effective change. *Canadian*

*psychology/Psychologie canadienne* 49, 3 (2008), 186.

[101] Richard M. Ryan, Martin F. Lynch, Maarten Vansteenkiste, and Edward L. Deci. 2011. Motivation and Autonomy in Counseling, Psychotherapy, and Behavior Change: A Look at Theory and Practice. *The Counseling Psychologist* 39, 2 (2011), 193–260. https://doi.org/10.1177/0011000009359313 arXiv:https://doi.org/10.1177/0011000009359313

[102] Michael J. Scheel. 2011. Client Common Factors Represented by Client Motivation and Autonomy. *The Counseling Psychologist* 39, 2 (2011), 276–285. https://doi.org/10.1177/0011000010375309 arXiv:https://doi.org/10.1177/0011000010375309

[103] Jessica Lee Schleider, Mallory Dobias, Jenna Sung, Emma Mumper, and Michael C Mullarkey. 2020. Acceptability and utility of an open-access, online single-session intervention platform for adolescent mental health. *JMIR mental health* 7, 6 (2020), e20513.

[104] Toni Schmader and Constantine Sedikides. 2018. State authenticity as fit to environment: The implications of social identity for fit, authenticity, and self-segregation. *Personality and Social Psychology Review* 22, 3 (2018), 228–259.

[105] Woosuk Seo, Chanmo Yang, and Young-Ho Kim. 2024. ChaCha: Leveraging Large Language Models to Prompt Children to Share Their Emotions about Personal Events. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 903, 20 pages. https://doi.org/10.1145/3613904.3642152

[106] Ashish Sharma, Kevin Rushton, Inna Lin, David Wadden, Khendra Lucas, Adam Miner, Theresa Nguyen, and Tim Althoff. 2023. Cognitive Reframing of Negative Thoughts through Human-Language Model Interaction. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Anna Rogers, Jordan Boyd-Graber, and Naoaki Okazaki (Eds.). Association for Computational Linguistics, Toronto, Canada, 9977–10000. https://doi.org/10.18653/v1/2023.acl-long.555

[107] Ashish Sharma, Kevin Rushton, Inna Wanyin Lin, Theresa Nguyen, and Tim Althoff. 2024. Facilitating self-guided mental health interventions through human-language model interaction: A case study of cognitive restructuring. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–29.

[108] Jonathan Smallwood. 2013. Distinguishing how from why the mind wanders: a process–occurrence framework for self-generated mental activity. *Psychological bulletin* 139, 3 (2013), 519.

[109] Ryan Smith, Samuel Taylor, Robert C Wilson, Anne E Chuning, Michelle R Persich, Siyu Wang, and William DS Killgore. 2022. Lower levels of directed exploration and reflective thinking are associated with greater anxiety and depression. *Frontiers in Psychiatry* 12 (2022), 782136.

[110] Joshua M Smyth, Christopher F Anderson, Jill R Hockemeyer, and Arthur A Stone. 2002. Does emotional non-expressiveness or avoidance interfere with writing about stressful life events? An analysis in patients with chronic illness. *Psychology and Health* 17, 5 (2002), 561–569.

[111] Joshua M Smyth and James W Pennebaker. 2008. Exploring the boundary conditions of expressive writing: In search of the right recipe. *British journal of health psychology* 13, 1 (2008).

[112] C Rick Snyder, Susie C Sympson, Florence C Ybasco, Tyrone F Borders, Michael A Babyak, and Raymond L Higgins. 1996. Development and validation of the State Hope Scale. *Journal of personality and social psychology* 70, 2 (1996), 321.

[113] Inhwa Song, Sachin R Pendse, Neha Kumar, and Munmun De Choudhury. 2024. The typing cure: Experiences with large language model chatbots for mental health support. *arXiv preprint arXiv:2401.14362* (2024).

[114] Yan Tao, Olga Viberg, Ryan S Baker, and René F Kizilcec. 2024. Cultural bias and cultural alignment of large language models. *PNAS nexus* 3, 9 (2024), pgae346.

[115] Gemini Team, Rohan Anil, Sebastian Borgeaud, Yonghui Wu, Jean-Baptiste Alayrac, Jiahui Yu, Radu Soricut, Johan Schalkwyk, Andrew M Dai, Anja Hauth, et al. 2023. Gemini: a family of highly capable multimodal models. *arXiv preprint arXiv:2312.11805* (2023).

[116] Kate Thompson. 2004. Journal writing as a therapeutic tool. In *Writing cures*. Routledge, 90–102.

[117] Gabriele Travagin, Davide Margola, and Tracey A Revenson. 2015. How effective are expressive writing interventions for adolescents? A meta-analytic review. *Clinical psychology review* 36 (2015), 42–55.

[118] Liuping Wang, Xiangmin Fan, Feng Tian, Lingjia Deng, Shuai Ma, Jin Huang, and Hongan Wang. 2018. mirrorU: scaffolding emotional reflection via in-situ assessment and interactive feedback. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–6.

[119] Jing Wei, Sungdong Kim, Hyunhoon Jung, and Young-Ho Kim. 2024. Leveraging Large Language Models to Power Chatbots for Collecting User Self-Reported Data. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW1, Article 87 (apr 2024), 35 pages. https://doi.org/10.1145/3637364

[120] Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V Le, Denny Zhou, et al. 2022. Chain-of-thought Prompting Elicits Reasoning in Large Language Models. In *Advances in Neural Information Processing Systems (NeurIPS '22)*. 24824–24837.

[121] Yuewen Yang, Thalia Viranda, Anna R. Van Meter, Tanzeem Choudhury, and Daniel A. Adler. 2024. Exploring Opportunities to Augment Psychotherapy with Language Models. In *Extended Abstracts of the 2024 CHI Conference on Human Factors in Computing Systems (CHI EA '24)*. Association for Computing Machinery, New York, NY, USA, Article 144, 8 pages. https://doi.org/10.1145/3613905.3650990

[122] Chunpeng Zhai, Santoso Wibowo, and Lily D. Li. 2024. The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review. *Smart Learning Environments* 11, 1 (18 Jun 2024), 28. https://doi.org/10.1186/s40561-024-00316-7

[123] Mingqian Zheng, Jiaxin Pei, Lajanugen Logeswaran, Moontae Lee, and David Jurgens. 2024. When "A Helpful Assistant" Is Not Really Helpful: Personas in System Prompts Do Not Improve Performances of Large Language Models. arXiv:2311.10054 [cs.CL] https://arxiv.org/abs/2311.10054

[124] Minjun Zhu, Linyi Yang, and Yue Zhang. 2024. Personality Alignment of Large Language Models. arXiv:2408.11779 [cs.CL] https://arxiv.org/abs/2408.11779

# A   Formative Study Material

## A.1   Example Narrative

*Ever since middle school, I've always seen myself as someone who thrives in the shadows, away from the glaring spotlight of public attention. I preferred to keep my head down, diligently working in quiet solitude, and my social interactions were always subtle, never the center of attention. But recently, my usual tranquility at work was disrupted when I was unexpectedly chosen to lead a major project. This wasn't just any project; it involved a division-wide presentation, thrusting me into the very spotlight I had always avoided. The thought of standing in front of all those people, having to lead and direct, filled me with an intense dread that was completely new to me.*

# B   Summarized Task Instructions for Each Generation Pipeline

## B.1   Theme Generation

- Role: You are a therapeutic assistant designed to foster deep self-reflection and promote personal growth in clients. Your approach is empathetic, client-centered, and rooted in the principles of therapeutic inquiry.
- Input type and format:
  `<initial_information/>`: Client's initial brief introductory of difficulty narrative, and the client's background.
  `<previous_session_log>`: Logs of sessions before the current session. DO NOT overlap with the previously selected themes.
- Output:
  - main_theme: Each theme from the user's initial narrative and previous log. Align closely with the user's language, expressions.
  - expressions: An array of diverse different expressions of the main_theme. When appropriate, introduce metaphoric expressions or nuanced language that might provide additional therapeutic value, but always anchor these in the client's original phrasing and emotional context.
  - quote: Most relevant part of the user's log to the theme

## B.2   Question Generation

- Role: You are a therapeutic assistant specializing in generating socratic questions to facilitate self-reflection and personal growth in clients. Per each session within the system, the client brings up a Theme in one's narrative that one would like to navigate about.
- Task: Given a client's personal narrative and context, your task is to generate list of Socratic questions and intention of the question
- Input type and format:
  `<initial_information/>`: Client's initial brief introductory of difficulty, and the client's background.
  `<previous_session_log>`: Logs of sessions before the current session.
  `<theme_of_session/>`: Theme of the current session.

## B.3   Keyword Generation

- Role: You are a therapeutic assistant specializing in supporting users to response to Socratic questions, facilitating self-reflection and personal growth.
- Task: The user is given a Socratic question to think about. However, it's not cognitively easy to answer those questions. Therefore, your task is to provide 'keywords' or 'short phrases' that might be useful for user to answer the given question. These keywords might act as 1) cognitive scaffolding 2) activate what might have been blind spot of the user 3) what might be relevant to users context 4) and so on.
- Input type and format:
  `<initial_information/>`: Client's initial brief introductory of difficulty, and the client's background.
  `<previous_session_log>`: Logs of sessions before the current session.
  `<question/>`: Question that the user is trying to response now.

## B.4   Comments Generation

- Role: You are a therapeutic assistant specializing in supporting users to work on personalized self-help workbook.
- Task: While the user is responding to the Socratic question, your task is to provide useful comments based on the context of user's response. For example, if the user hasn't started answering the question or is in very early phase, you might a) provide tips on how to approach the question. You might also provide b) encouraging, affirmation, or supportive feedback in a user-tailored manner. It also could be the c) sub-question in answering the given question, or some d) insightful comment, or e) any other type of comment that might support user in responding to the question.
- Input type and format:
  `<initial_information/>`: Client's initial brief introductory of difficulty, and the client's background.
  `<previous_session_log>`: Previous Logs of the session before the current session.
  `<question/>`: Question that the user is trying to think about now.
  `<current_response>`: Current response status of the user

## B.5  Summary Generation

- Role: You are a therapeutic assistant helping the user reflect on their session and progress.
- Task: Summarize the user's experiences and insights from their Q&A log into a coherent and concise narrative. Focus on the essence of their reflections without overemphasizing any one aspect.
- Guidelines:
  - Capture the key points and overall sentiment without unnecessary detail.
  - Use the user's own language and expressions where appropriate.
  - Keep the summary realistic, proportional to the content of the user's log, and grounded to user's input.
  - Feel free to draw on the following as needed
    * Major themes or emotions
    * Notable progress or changes in perspective
    * Encouragement to continue reflecting and growing
    * Recognition of the user's strengths and resources

## C  Statements of Pathway Subscale of Adult State Hope Scale

The validated Korean version [16] of the Pathway subscale of the Adult State Hope Scale [112] consists of the following four questions:

(1) If I should find myself in a jam, I could think of many ways to get out of it.
(2) There are lots of ways around any problem that I am facing now.
(3) I can think of many ways to reach my current goals.
(4) I know that even when others despair, I can find a way to solve my problems.

Each question is rated on an 8-point rating scale, with scores and labels as in the following:

(1) Definitely False
(2) Mostly False
(3) Somewhat False
(4) Slightly False
(5) Slightly True
(6) Somewhat True
(7) Mostly True
(8) Definitely True