



# Chorus of the Past: Toward Designing a Multi-agent Conversational Reminiscence System with Digital Artifacts for Older Adults

Jingwei Sun  
Lenovo Research  
Beijing, China  
sunjw12@lenovo.com

Nianlong Li  
Lenovo Research  
Beijing, China  
linl1@lenovo.com

Liuxin Zhang  
Lenovo Research  
Beijing, China  
zhanglx2@lenovo.com

Zhongyue Zhang  
The Hong Kong University of Science  
and Technology (Guangzhou)  
Guangzhou, China  
zzhang837@connect.hkust-gz.edu.cn

Zhangwei Lu  
Lenovo Research  
Beijing, China  
luzw3@lenovo.com

Yu Zhang\*  
Southeast University  
Nanjing, China  
Lenovo Research  
Beijing, China  
zhangyu29@lenovo.com

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

Mengyang Wang  
The Hong Kong University of Science  
and Technology (Guangzhou)  
Guangzhou, China  
mwang492@connect.hkust-gz.edu.cn

Yan Xiang  
The Hong Kong University of Science  
and Technology (Guangzhou)  
Guangzhou, China  
yxiang411@connect.hkust-gz.edu.cn

Qianying Wang\*  
Lenovo Research  
Beijing, China  
wangqya@lenovo.com

## Abstract

Reminiscence has been shown to provide benefits for older adults, but traditionally relies on personal photos as memory cues and interactions with real people who may not always be available. We present ReminiBuddy, a novel LLM-powered multi-agent conversational system, which allows older adults to engage with two distinct agents—one embodying an older identity and the other a younger identity—while using not only personal photos but also 3D models of generic nostalgic objects as memory cues. Our study, with older adult participants, found that the conversational approach both enjoyable and beneficial for reminiscence. While the younger agent was perceived as more emotionally engaging, the older one fostered

greater resonance in content. Personal photos prompted autobiographical memories, whereas 3D generic nostalgic objects evoked shared memories of an era, contributing to a more multifaceted reminiscence experience. We further present design implications for better supporting older adults in reminiscing with LLM-powered conversational agents.

## CCS Concepts

- Human-centered computing → Accessibility design and evaluation methods; User studies.

## Keywords

Multi-agent Conversational System, Conversational agents, Glasses-free 3D Monitor, Reminiscence, Human-AI Interaction, Older Adults

\*Corresponding authors.



This work is licensed under a Creative Commons Attribution 4.0 International License.  
*CHI '25, Yokohama, Japan*  
© 2025 Copyright held by the owner/author(s).  
ACM ISBN 979-8-4007-1394-1/25/04  
<https://doi.org/10.1145/3706598.3713810>

## ACM Reference Format:

Jingwei Sun, Zhongyue Zhang, Mengyang Wang, Nianlong Li, Zhangwei Lu, Yan Xiang, Liuxin Zhang, Yu Zhang, Qianying Wang, and Mingming Fan. 2025. Chorus of the Past: Toward Designing a Multi-agent Conversational Reminiscence System with Digital Artifacts for Older Adults. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April*

26–May 01, 2025, Yokohama, Japan. ACM, New York, NY, USA, 22 pages.  
<https://doi.org/10.1145/3706598.3713810>

## 1 Introduction

Reminiscence, the process of revisiting past memories, is a well-established practice known to offer significant benefits for older adults, including improvements in mental health, mood, and social interactions [35, 43, 44]. Engaging in discussions about past experiences serves as a cognitively stimulating activity for older adults, and contributes to psychological well-being through life review [8, 13, 15]. This process is complex and multifaceted, varying significantly depending on the context and individuals involved. It can strengthen identity or serve as a means to educate and inform others [41]. The specific functions of reminiscence often depend on the nature of the relationship between participants. For example, sharing memories with close family members typically emphasizes identity reinforcement, whereas interactions with younger generations or strangers may focus more on teaching or advice-giving [15].

Despite the well-documented benefits, older adults often encounter challenges when engaging in reminiscence. Barriers such as difficulty in recalling memories, lack of opportunities for meaningful conversation, or the absence of interested and available conversation partners can diminish the frequency and depth of reminiscence, thereby limiting its positive effects [38]. In response, the HCI community has endeavored to design technologies that not only facilitate but also enhance the practice of reminiscence by creating social scenarios that promote storytelling with families [28, 32] and memory sharing among older adults [7].

While prior technologies have provided valuable tools to support reminiscence, they often rely on interactions with real people—be it family members, caregivers, or even strangers [7, 28, 32]. This reliance on human interlocutors can sometimes limit the availability and accessibility of reminiscence opportunities. To address these limitations, recent research has begun exploring the use of conversational AI-agent-driven systems designed to simulate human-like dialogue—as a novel approach to supporting reminiscence [9]. Conversational agents powered by Large Language Models (LLMs) trained on vast datasets, can understand and generate human-like text, enabling these agents to engage in natural, meaningful conversations [21, 22]. These agents can simulate the role of a conversational partner by asking relevant questions, providing empathetic responses, and recalling previously shared information, thereby creating a coherent and personalized dialogue experience [18, 39]. This shift from human to AI interaction introduces new possibilities for reminiscence, offering older adults a more accessible and consistent means of recalling and sharing their memories, even in the absence of human partners.

On the other hand, memory cues are also important for enhancing the reminiscence experience. Visual cues, particularly photos, are among the most effective triggers for reminiscence, as autobiographical memories are often encoded visually [12, 19]. For older adults, storytelling around photos not only evokes memories but also fosters social bonds, especially in intergenerational settings. This process contributes to the creation of a legacy through the

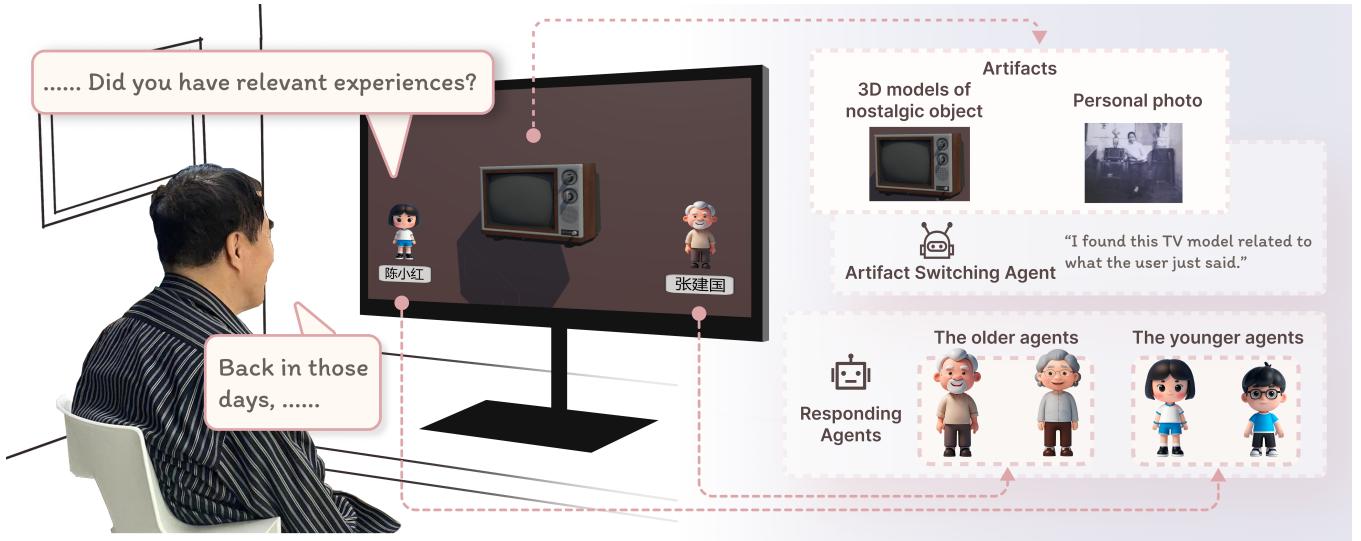
sharing of personal mementos [30]. While personal photos are commonly used in reminiscence practices, research has shown that generic images can elicit more emotionally rich and personal stories [2], offering a different and potentially complementary role to that of personal photos.

HCI researchers have investigated technologies to provide meaningful cues [36] to support reminiscence. Yet, previous practices typically involved disjointed memory cues [9], or relied on limited dimensions (e.g., location [36]) or user selection and annotations [4] to create a meaningful reminiscence experience, without seamlessly weaving these memory cues into the process. With the advent of multimodal LLMs, the capability to comprehend data beyond text has greatly improved [46]. These multimodal LLM-powered systems can generate coherent and vivid narratives based on visual inputs [47], and even produce extended multimodal stories using user-provided images and texts as initial prompts [45]. This advancement opens new opportunities for conversational agents to engage users in dialogue centered around diverse modalities of memory cues, potentially enriching the reminiscence experience.

Despite these technical advances, there remains a significant gap in understanding how these user-agent conversational systems can be effectively designed and integrated into the practice of reminiscence. Specifically, the interactive dynamics between users and agents, the potential benefits of interacting with AI agents, and the utilization of various types of memory cues, are areas that require further exploration. In light of these research gaps, we introduced a multi-agent conversational system, ReminiBuddy, which is designed to explore the impact of different agent identities and nostalgic artifacts on reminiscence among older adults (Figure 1). When interacting with ReminiBuddy, older adults engage in simultaneous conversations with two LLM-powered agents: one embodying an older identity and the other a younger identity, offering diverse perspectives. They engage in conversation around nostalgic artifacts, including personal photos and 3D models of generic nostalgic objects, creating a richly interactive environment for memory recall. Our system served as a technology probe to investigate the following **research questions**:

**How should such LLM-powered conversational system be designed, and in what ways might such a system influence reminiscence behavior, in terms of: (A) the conversation dynamics; (B) agent design; (C) digital artifacts as memory cues.**

We conducted a user study with older adult participants, in which they engaged in conversations freely with agents of varying identities for 30–40 minutes, discussing 3D generic nostalgic objects and personal photos displayed on the screen. Results indicated that participants found the conversational process both enjoyable and beneficial for reminiscence. Regarding agent identities, while participants perceived the younger agents as more emotionally engaging, they felt that the older agents fostered greater resonance in content. As for artifacts, personal photos elicited more autobiographical memories, whereas 3D models of generic nostalgic objects evoked shared memories of an era, prompting reflection on the present. The combination of these diverse elements contributed to a richer, more multifaceted reminiscence experience. We have further discussed the implications of these findings for the design of multi-agent



**Figure 1:** ReminiBuddy prototype. An older adult interacts with two *Responding Agents* with two identities (older and younger) on a glasses-free 3D monitor, discussing nostalgic artifacts (3D models of generic nostalgic objects or personal photos). The older adult can freely choose which agent to converse with. An *Artifact Switching Agent* determines which artifact to display based on conversation history for each round of conversation.

systems to support reminiscence among older adults, and we highlight new directions to explore for fostering more engaging and meaningful reminiscence experiences.

In conclusion, this work makes the following contributions:

- We designed an LLM-powered multi-agent conversational reminiscence system, ReminiBuddy, which engages older adults in reminiscence with two-identity agents through discussions of personal photos and 3D models of generic nostalgic objects.
- We conducted a user study to understand how the different agent identities and nostalgic artifacts facilitate reminiscence among older adults. We further present design implications for better supporting older adults in reminiscing with LLM-powered conversational agents.

## 2 Related Work

### 2.1 Reminiscence and Older Adults

Reminiscence, the process of revisiting past memories, offers significant benefits for older adults, including enhancements in mental health, mood, and social interactions [35, 43, 44]. Engaging in conversations about the past is a central activity in reminiscence, fostering a stronger sense of identity and serving as a cognitively stimulating exercise [13, 15]. Such engagement contributes to psychological well-being by facilitating life review [8]. The process of reminiscence is inherently complex, serving multiple functions that address the emotional, social, and cognitive needs of older adults. These functions include problem-solving, identity reinforcement, and teaching or informing others [41]. Research indicates that in daily life, the functions of reminiscence vary depending on the social context and the relationship with the conversation partner [15]. For example, reminiscence with close partners or children tends to

emphasize identity functions, such as expressing self-continuity, where a well-known listener encourages elaboration. Conversely, with children or strangers, the teaching or informing function is more prominent, where the speaker shares past experiences to impart knowledge or offer advice.

Reminiscence is often triggered by various cues, including verbal prompts, visual stimuli (e.g., photos), music, and scents. Among these, visual cues are particularly effective in eliciting reminiscence, as autobiographical memories are frequently encoded as visual images [12, 19]. In everyday life, older adults find storytelling with photos to be both enjoyable and socially connecting, as it facilitates intergenerational relationships through the creation of artifacts like mementos, thus enabling them to leave a lasting legacy [30]. The use of photos as triggers for reminiscence may produce different effects depending on whether the photos are personal or generic. While personal photos are commonly employed in reminiscence therapy to evoke autobiographical memories (e.g., family photos), a study by Astell et al. [2] demonstrated that individuals with dementia recounted more stories with emotional and personal significance when shown generic photos compared to personal ones.

### 2.2 Challenges and Technologies to Support Reminiscence of Older Adults

While reminiscence offers significant benefits to older adults, its practice is often hampered by challenges in daily life. A critical issue is that older adults, despite enjoying storytelling and sharing memories, frequently feel that they have nothing to share or that no one is interested in listening [38]. This perception can severely limit the availability of materials and opportunities for conversations, thereby hindering their engagement in reminiscence activities and depriving them of potential psychological and social benefits.

Over the past few decades, the HCI community has increasingly focused on designing technologies to support and enhance the reminiscence practices of older adults. These efforts have largely centered on addressing the challenges of providing meaningful memory cues and materials, as well as creating scenarios that encourage storytelling and memory sharing around personal artifacts, such as photos. Early initiatives, such as Memory Karaoke [36], retrieved the location, time, and photos from older adults' personal visits to evoke meaningful memories, thereby facilitating reflection on their past experiences. More recent work, like MomentMeld [23], leverages AI technology to align semantically similar personal photos across generations, fostering intergenerational interactions and empathy. Other research has explored organizing photos for more effective family sharing through storytelling, as demonstrated by Axtell et al. [3]. Emerging technologies, including Virtual Reality (VR) and Augmented Reality (AR), have also shown potential in enabling remote collective reminiscence and augmenting 3D artifacts for intergenerational storytelling [7, 28].

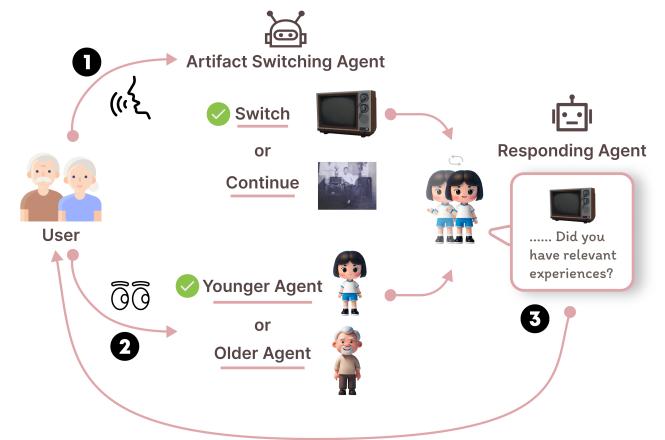
While existing reminiscence technologies support older adults in recalling memories, whether independently or through interactions with both familiar people and strangers, they all involve real human interactions. In contrast, conversational agents offer a distinct experience by enabling older adults to engage in memory recall and storytelling through dialogue with an AI rather than a real person. This shift presents a unique opportunity to explore how these agents can facilitate reminiscence differently from traditional social interactions. However, the dynamics of reminiscing with an agent, and the potential benefits of this approach, remain underexplored, revealing a critical gap in current research. Much of the existing research has focused on the technical feasibility of conversational agents and the development of content-generation technologies. For example, Caros et al. [9] developed text-based chatbot agents powered by AI to simulate conversational scenarios. Assistive robots equipped with dialogue systems and knowledge of older adults' photos have also shown promise in facilitating reminiscence while providing cognitive training [39]. More recently, the rise of LLMs has enabled the generation of natural conversations, as demonstrated by Jeung et al. [21], who explored the potential of LLMs to support memory recall through AI-generated questions and cues. However, these studies primarily focused on content generation and technical implementation, with less attention paid to the interactive dynamics and the quality of user-agent engagement during reminiscence activities.

In addition to their conversational capabilities, LLMs offer significant potential for multimodal understanding, which can enhance reminiscence experiences through the integration of both visual and textual cues. For example, the AI-driven interactive photo album GoodTimes has demonstrated the benefits of using multimodal inputs like photos and narratives to engage users in personalized storytelling, creating deeper emotional connections during reminiscence [40]. The ability to process and understand multiple forms of media allows for a richer interaction experience, helping older adults to more effectively recall memories and engage with personal artifacts in ways that traditional methods or singular media cues may not support.

Building on these foundations, our research introduces a multi-agent conversational system designed specifically to support reminiscence among older adults. This system integrates both personal photos and 3D models of generic nostalgic objects, offering a rich, interactive environment for users. We analyze older adults' preferences for different agents within the system, as well as their interactions with these agents and the broader system. By shifting the focus from content generation to understanding user preferences and interaction patterns, our work aims to provide more meaningful and emotionally resonant reminiscence experiences for older adults, thereby maximizing the benefits of these activities.

### 3 ReminiBuddy: Design and Implementation

Based on prior work and inspirations, we designed a multi-agent conversational reminiscence system, ReminiBuddy, with digital artifacts for older adults to engage in conversations with the agents. As Figure 1 shows, the system employs a glasses-free 3D monitor, displaying a nostalgic artifact—either a generic nostalgic object 3D model or a personal photo—in the center, with two agents—an older agent and a younger agent—positioned in the bottom left and right corners respectively. The participant sits in front of the 3D monitor and chats with the two agents about the central artifact through voice interaction. The system's overall interaction flow is illustrated in Figure 2. In this section, we described the interface and interaction features of the system, as well as its implementation details.



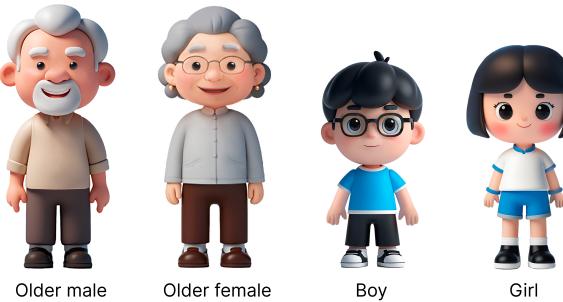
**Figure 2: Interaction flow of the ReminiBuddy system.** (1) After the user speaks, the *Artifact Switching Agent* determines whether to switch to a new artifact or continue with the current one for the next conversation round. (2) Simultaneously, the system selects the *Responding Agent* (younger or older) based on the user's eye gaze during speech. (3) Finally, the chosen *Responding Agent* generates and vocalizes a response to the user.

#### 3.1 Agent Design

**Feature 1: An Older Agent and a Younger Agent as the Responding Agents.** Research on reminiscence among older adults

has primarily focused on two contexts: peer communication [7] and inter-generational storytelling [28, 30, 42]. While peer communication facilitates better social support. By incorporating agents with varied characteristics and conversational styles, we could encourage older adults to adapt their storytelling style and content, addressing the multifaceted nature of reminiscence.

We implemented two types of *Responding Agents*: an older agent and a younger agent. The older agent is characterized as a Chinese older adult, approximately 65-75 years old, designed to be peers with our participants. The younger agent is depicted as a Chinese child of about 12 years old. The appearance of these avatars is designed to align with the Chinese cultural context, mirroring our participants' cultural background, enhances participants' sense of connection through cultural resonance [29]. For each role, we designed two avatars (one male and one female) in an animated style (Figure 3), assigning them common names and voices that correspond to their role settings.



**Figure 3: The four *Responding Agents*: two older agents (male and female) and two younger agents (boy and girl).**

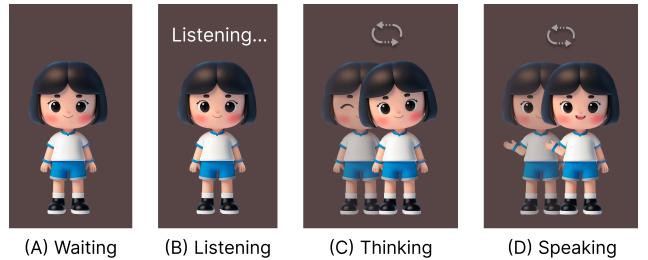
Regarding the conversation design of agents, in each round of conversation, the agents would first offer either content-based or emotional responses to what the participant said. Then they would share personal experiences or knowledge related to the currently displayed artifact, and encourage the participant to share their own stories and feelings. Following are two examples of the agents' responses:

*"Oh wow, the way you put it, I can totally picture that scene in my head! Speaking of music, I notice there's another old gem here - a radio recorder (illustrated in Figure 5C). Take a look at this big black device, with those large speakers on both sides and all those buttons and knobs in the middle. Isn't this the very thing we used to play the 'Huo Yuanjia' theme song on? Do you remember what it was like listening to music on one of these? I recall how everyone loved using it to record radio shows and their favorite songs. What kind of music or radio programs did you enjoy listening to most back then?" - An older agent response, with a radio-cassette recorder model as the displayed artifact*

*"Wow, nainai (a generic Chinese term for older female), you were so brave! Standing up to your father for the hairstyle you wanted - that reminds me of when I wanted*

*to dye my hair and my mom didn't agree. You must have felt so cool insisting on getting that perm back then! I really admire your courage. And you mentioned it was your first time going to a photo studio alone - you must have been nervous and excited at the same time! I bet you were quite independent and strong-willed when you were young. I wonder, when you see all the trendy hairstyles young people have nowadays, do you find them interesting?" - A younger agent response, with a personal photo as the displayed artifact*

To help participants better understand the current state of the system during the conversation, especially during the process of waiting for the agents to generate responses, we designed four states of each agent (Figure 4): *waiting* for the user to respond, *listening* to user input, *thinking* about how to respond, and *speaking* to the user. In each conversational round, one of the *Responding Agents* (the logic for determining which agent responds is detailed in the following section) generates a response to the user or initiates the conversation at the outset.



**Figure 4: Four states of *Responding Agents* illustrated by the girl agent.**

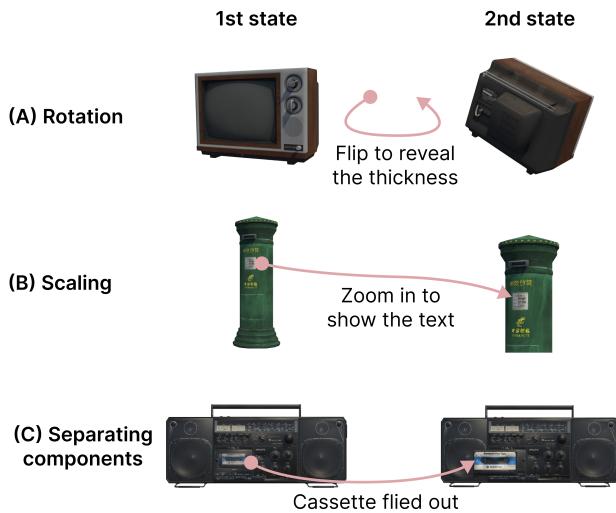
### 3.2 Digital Artifacts as Memory Cues

**Feature 2: 3D Models of Generic Nostalgic Objects and User's Personal Photos as Memory Cues.** Previous research on reminiscence among older adults has often utilized personal photos as stimuli, encouraging older adults to engage in storytelling, authoring personal histories, and meaning-making [5, 37]. However, personal photos capture only a fraction of an older adult's memories, leaving many experiences undocumented. Research has shown that generic images can sometimes play a different role of personal photos, and evoke more emotionally rich and personal stories [2].

Based on this, we incorporated generic nostalgic objects and user's personal photos as memory cues to enrich and complement the reminiscence content. Generic nostalgic objects refer to items that are no longer commonly seen or have undergone significant design changes over time, but were ubiquitous during the participants' youth. We selected objects that were common or held particular importance to the entire generation when our participants were young, approximately 20-50 years ago.

In addition, research has shown that stimuli in immersive formats, such as 3D models, could enrich storytelling [28]. Consequently, integrating various forms of stimuli into the system could provide a more comprehensive and engaging reminiscence experience. Therefore, we employed 3D model representations for the

generic nostalgic objects, aiming for a more immersive and enriched storytelling experience. For each 3D model of generic nostalgic objects, we incorporated two states (Figure 5). The first presents a conventional view, offering an overall representation of the object. The second state is designed to highlight a specific feature or reveal a less obvious aspect of the item, achieved through rotation, scaling, or separating components. For instance, the TV model's second state rotates to reveal its characteristically thickness (Figure 5A). For the mailbox model, the second state zooms in to show the text on it more clearly (Figure 5B). For the cassette-radio recorder, the second state shows the cassette flying out of the device, displaying closer to the user (Figure 5C). These model variations are intended to diversify the user's points of interest when examining the objects.



**Figure 5: Illustrations of the two states of generic nostalgic objects.**

### 3.3 Conversation Dynamics

**Feature 3: Employ an Artifact Switching Agent to Manage Conversational Topic.** To manage the progression of conversational topics, we implemented an *Artifact Switching Agent*. This agent determines whether to introduce a new artifact for the next round of conversation and, if so, which artifact to display. It operates by searching the artifact database for items semantically related to the user's most recent utterance. The decision-making process considers several factors, including the user's latest utterance, perceived user interest in the current topic, the number of conversational rounds regarding the current artifact, and the types of artifacts (generic nostalgic objects or personal photos) previously displayed to ensure a diverse selection. This approach, that is to separate the *Artifact Switching Agent* and the *Responding Agents*, is inspired by research suggesting that multi-step prompting techniques can enhance the effectiveness of LLM outputs, thereby reducing unexpected responses [20, 25]. During prototype development, we experimented with using a single LLM request to both select the artifact for discussion and generate the agent's response.

However, this approach often led to irrelevant outputs, such as incorporating the artifact image's filename into the conversation (e.g. “*This photo ‘xxx.jpg’ depicts an old television...*”) or mentioning details of the image retrieval process (e.g. “*You might be interested in this photo from the list.*”), which disrupted the conversational flow and confused the user. At the system's initiation, a random artifact—either a generic nostalgic object or a personal photo—is selected from the database. The *Artifact Switching Agent*'s decisions are then communicated to the *Responding Agents*, guiding their subsequent conversations. Notably, the user is unaware of the *Artifact Switching Agent*'s existence, maintaining a natural conversational experience.

**Feature 4: Use Voice Interaction to Navigate the Conversation Flow.** Given the cognitive decay often experienced by older adults, their adoption and learning curve of digital technologies typically lag behind those of younger individuals [14]. To reduce the cognitive load on older adults while using our system, the interaction mode should be designed to simulate natural interaction as closely as possible.

We adopted voice interaction for both input and output to achieve a more natural interface, minimizing the need for reading or manual input. When a user decides to speak, they initiate their response by pressing the spacebar on the keyboard, which triggers the agents to change to the *listening* state. The user then presses the spacebar again to signal the end of their voice input. While this method may not be the most intuitive for voice interaction, we believe it represents the optimal approach given current technological constraints and the typical behavior patterns of older adult users. Specifically, older adults often require additional time to formulate their thoughts [24]. The spacebar initiation provides a pause for reflection without inducing pressure to respond immediately. Moreover, older adults tend to speak at a slower pace, posing a challenge for the system in accurately determining when the user has finished speaking. Therefore, using a key pressing indicator is preferable.

Each agent is aware of the conversation history between the other agent and the user, but the two agents do not converse with each other. In addition, we implemented an interruption feature. Users can interrupt an agent's speech at any point by pressing the spacebar and beginning to speak, which immediately halts the agent's vocalization.

**Feature 5: Determine Responding Agent Based on the User's Eye Gaze.** The system determines which agent will respond based on the user's eye gaze during his/her speaking turn. The agent receiving the longer gaze duration responds; if the gaze duration difference is less than 1 second, a random agent is selected to reply. This feature was employed to mimic the natural interaction in daily conversation, helping the older adults focus on the conversations rather than on operating the system.

### 3.4 Implementation

The system is implemented in Unity. It listens and understands the user's vocal response, determines which agent should respond, decides whether to change the displayed artifact for the ensuing conversation, and generates the agent's response and delivers it through voice.

**3.4.1 User Input.** The user's vocal response is recorded and transcribed into text using the CT-Transformer model [10]. Eye gaze is tracked using Tobii Eye Tracker 5L.

**3.4.2 Artifact Switching Decision.** Personal photos and generic nostalgic objects (each object featuring two states) were stored into the database in advance. Besides, we used Claude 3.5-Sonnet to generate textual descriptions for these personal photos and object states. The details of the prompt design are discussed in Section 3.4.4. These descriptions were converted into embeddings and saved in the database.

The artifact decision process is illustrated in Figure 6. During each round of conversation, if the currently displayed artifact is the first state of a model, its corresponding second state is retrieved. Otherwise, the system retrieves the top-5 most similar artifact descriptions from the database based on the user's latest utterance using vector cosine similarity search while filtering out previously shown artifacts. The *Artifact Switching Agent* then decides, based on the conversation history between the user and the agents, whether to switch to a new artifact (or a new state of the model) for display or to maintain the current one. If a switch is deemed appropriate, the agent also determines which artifact to display next. This decision is then communicated to the next *Responding Agent* to inform its response generation.

**3.4.3 Agent Response Generation.** Agents' responses are generated by Claude-3.5 Sonnet, with prompts listed in Appendix C. The generated responses are converted into speech using Azure AI Speech. The voices for the four agents are as follows: "Yunze" with "SeniorMale" speaking style for the older male agent, "Xiaoqiu" with "Default" speaking style for the older female agent, "Yunxia" with "Default" speaking style for the younger male agent, and "Xiaoyi" with "Default" speaking style for the younger female agent.

**3.4.4 Prompt Engineering.** We adopted an iterative approach to prompt engineering, refining prompts for artifact descriptions, artifact switching decisions, and responding agents. Five adult participants aged 25–40 participated in testing, focusing on the coherence of conversations and the relevance of selected artifacts in simulated reminiscence dialogues. Through multiple iterations, participants confirmed that the conversational flow and artifact selection were appropriate. Subsequently, a pilot study with an older adult participant was conducted in two separate sessions to evaluate the finalized prototype. The participant reported a positive reminiscence experience, and any issues identified during the pilot study were addressed.

Below, we provide detailed information on each aspect of prompt engineering:

In **describing artifacts** (with prompts and examples provided in Appendix A), we prioritize the most relevant details while ensuring the description flows naturally within the context of the conversation. Through iterative prompting and analysis of the model's responses, we found that achieving richer and more coherent descriptions required the LLM to capture not only the visible elements of the artifact but also the subtle atmosphere and broader context. Simple prompts like "describe this photo" often led to inconsistencies. To address this, we specified mandatory elements to include in descriptions for both personal photos and nostalgic objects, such

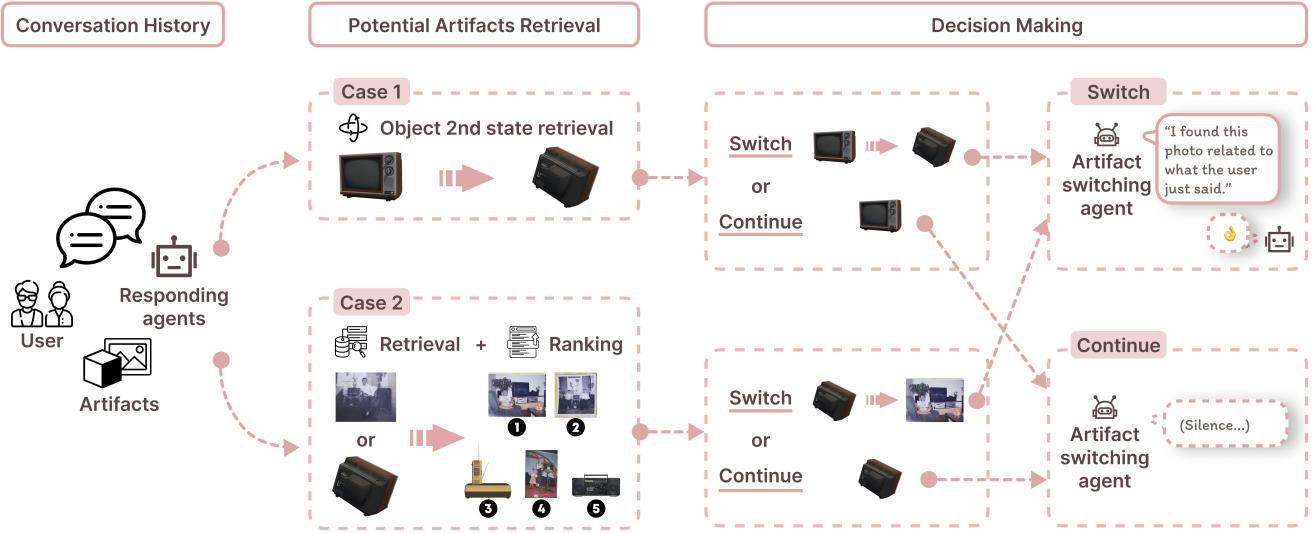
as people, facial expressions, object appearances, and relevant historical context. To enhance flexibility, we also prompted the LLM to suggest potential keywords that could emerge during reminiscence. This approach ensures the model generates comprehensive descriptions while still allowing room for the organic discovery of additional details. Testing showed that this strategy allowed the retrieval process to identify relevant objects, provide sufficient detail for ongoing discussion, and facilitate smooth topic transitions. No manual modifications were made to the artifact description after it was generated.

For the **prompts guiding the *Artifact Switching Agent* to determine the appropriate timing for introducing new, relevant artifacts into the conversation** (Appendix B), the goal is to enhance the older adults' experience by exposing them to a broader range of artifacts based on their interest level. We assume that if interest in a particular artifact wanes, the agent should switch to a new one. Through iterative testing, we developed expressions to gauge interest. For example, brief responses, limited personal anecdotes, or signs of disinterest indicate that it may be time to switch the artifact [6]. To further enrich the experience, we set a rule for the agent to introduce new artifacts after 4–5 conversational rounds, alternating between personal photos and other artifacts. Testing has shown that this strategy is effective in introducing new artifacts in a timely manner. However, we acknowledge that this approach may not always feel natural, as older adults may prefer to explore certain artifacts in greater depth if they find them engaging. Our method of gauging interest is primarily based on participant responses, and we discuss potential improvements—such as more context-aware approaches in the discussion section 6.1.3.

For **prompts that allow the *Artifact Switching Agent* to introduce new artifacts to the *Responding Agents*** (Appendix B)—essentially prompts that the responding agents receive, such as "switch the topic to this artifact"—we encountered several challenges. Directly presenting artifact descriptions to the responding agents often led to confusion. For example, personal photos might be mistakenly identified as not belonging to the older adults, and the context in which participants could view the objects was not always clear. In such cases, agents could inadvertently provide responses that were out of context for the older adults. Additionally, this approach sometimes resulted in repetitive questions or topics, especially when the same artifact was presented from a different angle in previous conversations. Transitions between artifacts also lacked smoothness. To address these challenges, we focused on crafting prompts that clearly communicated the context, ensuring that responding agents understood what participants were experiencing and could see during the interaction.

For the **role settings of the *Responding Agents*** (Appendix C), we kept the prompts simple, specifying the agents' ages, Chinese background, and ensuring they maintain a positive and encouraging attitude throughout the conversation.

For **prompts guiding the *Responding Agents* to generate responses** (Appendix C), iterative testing revealed a key issue: the agents were asking too many questions but failing to share personal experiences and feelings, which is uncommon in real-life reminiscence conversations. To address this, we adjusted the prompts to reduce the number of questions and focus more on sharing personal



**Figure 6: Artifact decision process.** The system handles two cases: if the current artifact is the first state of a 3D model, its second state is retrieved; Otherwise, the top-5 most similar artifacts are retrieved based on the user's latest utterance. The *Artifact Switching Agent* then decides whether to switch to a new artifact or continue with the current one, based on the conversation history.

experiences. We also introduced the *Artifact Switching Agent* condition, instructing the *Responding Agent* to follow topic guidance without direct interaction with the switching agent. During testing, we found that participants can occasionally ask questions, but the agent did not respond. We addressed this issue by incorporating appropriate responses into the context of the prompts.

## 4 User Study

### 4.1 Participants

Participants were recruited through word-of-mouth and snowball sampling methods. Eligibility criteria included adequate hearing ability and corrected vision, as well as the ability to engage in everyday conversations in Mandarin. Each participant received a 100 RMB compensation (equivalent to about 14 USD) for approximately 90 minutes of their time, which aligns with the local average pay for research participation. This study received approval from the Institutional Review Board.

A total of twelve participants (7 female, 5 male) completed the study. Their ages ranged from 60 to 80 (Mean = 66.58, SD = 5.42). None of the participants had prior experience with glasses-free 3D displays. Their demographic information is shown in Table 1.

### 4.2 Apparatus and Materials

In the user study, we ran our prototype on a Lenovo Thinkbook 16P laptop, which was connected to a Lenovo Thinkvision 27-inch glasses-free 3D monitor. The optimal viewing distance for the monitor is 0.6m – 1.2m.

For the generic nostalgic object 3D models, we prepared 9 models for each themed session: “home” and “going out.” The models for the “home” session include a television, radio-cassette recorder, early

mobile phone (brick phone), desk, rattan chair, refrigerator, sewing machine, gas cylinder, and ceiling fan. The models for the “going out” session include a camcorder, film camera, winter clothing set (hat, padded jacket, shoes), woven bag, bicycle, motorcycle, bus, public phone, and mailbox.

We asked each participant to prepare approximately 16 personal photos prior to the study, with eight for each theme. While the photos should ideally include elements related to our prepared nostalgic 3D objects, this is not mandatory. The photos were all taken before the year 2000.

### 4.3 Procedure

Participants were first informed of the aim of this study and signed a consent form. Experimenters then introduced the system’s key features and demonstrated their usage. Following eye-tracking calibration, participants engaged in a brief 5-minute practice session. This session displayed two 3D models sequentially, aimed at familiarizing participants with the system’s interface and interaction process.

Subsequently, participants engaged in two themed sessions: “home” and “going out,” presented in a randomized order. Each session lasted 15–20 minutes, depending on the interaction willingness of the participants, judged by the experimenter. The *Responding Agents’* genders were also randomized. In the first session, the agents’ genders were assigned randomly, and in the second session, their genders were set to be different from those in the first session (for both the older and younger agents), ensuring that participants could interact with all four agents. Agents’ positions (left or right) were also randomized.

**Table 1: Participants' demographic information.**

ID	Age	Sex	Daily digital device usage	Experience with conversational agents (e.g., smart speaker)	How often to organize or browse old photos
P1	63	M	3-5 hours	Yes	Never
P2	71	M	1-3 hours	No	Never
P3	67	F	3-5 hours	Yes	Rarely
P4	69	M	>5 hours	Yes	Sometimes
P5	63	F	1-3 hours	Yes	Rarely
P6	65	F	>5 hours	Yes	Sometimes
P7	80	M	1-3 hours	No	Often
P8	62	F	>5 hours	Yes	Rarely
P9	66	F	1-3 hours	Yes	Often
P10	70	F	3-5 hours	No	Sometimes
P11	60	M	3-5 hours	Yes	Never
P12	63	F	3-5 hours	No	Rarely

Upon completion of these two formal sessions, participants were asked to fill out a post-study survey using 5-point Likert scales (Appendix D). The survey gathered participants' experiences, focusing on the agents, artifacts, overall experience, any dizziness caused by the 3D display, and their use of technology and AI. For the conversational agents, we assessed factors such as likability, efficiency, trust, and the naturalness and smoothness of interactions, following established methods in prior HCI research on conversational agents [26, 50]. These factors were evaluated for each individual agent. Regarding artifacts, we examined their appropriateness, relevance, and usefulness in supporting reminiscence. Additionally, participants were asked about their willingness of using the system in daily life and whether it improved their mood, as an indicator of overall preference for the prototype. Subsequently, the experimenters conducted a semi-structured interview (outline listed in Appendix E) based on the survey results and observed usage patterns during the study, asking them to elaborate on their reasons behind their ratings and behaviors. Participants provided qualitative feedback on the system and shared their expectations for such systems.

#### 4.4 Data Collection and Analysis

**Conversation histories**, including roles, content, displayed artifacts, and timestamps, were recorded. Four authors collaboratively analyzed the conversation histories of three participants, employing both deductive and inductive coding techniques to develop a codebook for user response patterns and content. The four coders then independently coded the conversations, meeting subsequently to discuss and resolve disagreements. In the end, we arrived at three categories of user response patterns and seven categories of response content. For response content analysis, we quantified the occurrence of each category and word count (Chinese characters after transcription) as indicators of the reminiscence effect. To compare the effect of independent variables (e.g., agent identity or artifact type), we first weighted the dependent variable in each category to account for overall differences between independent variables for each participant, then applied Wilcoxon signed-rank tests to analyze the differences.

**Post-study survey** results were aggregated, and quartiles were calculated accordingly.

**Interviews** were recorded and transcribed. Three authors (also the study experimenters) analyzed the transcripts, identifying four themes: overall system perception, conversational dynamics, agent design, and artifact use. They then independently coded the transcripts, followed by meetings to discuss and resolve any coding discrepancies.

For data analysis, we examined participants' **ratings** from the post-study survey, including their overall experience as well as their perceptions of the agents and the displayed artifacts. We also analyzed the **conversation dynamics** with the agents, including speaking duration, number of speaking turns, and average speaking duration per turn. Additionally, we studied **content** of participants' utterances and explored the **effect of agent identity and artifact types** on conversation dynamics and response content. Finally, we investigated participants' **eye gaze patterns** during the interactions.

## 5 Results

In this section, we first present our key findings regarding participants' overall perceptions and attitudes towards the system. We then address our research question by examining participants' behaviors and attitudes in three aspects: conversation dynamics, agent design, and artifacts.

### 5.1 Overall System Perception and Attitudes

Overall, participants experienced an improvement in mood after engaging with the system (Mean = 4.3, Median = 4.0), as shown by the post-study survey. As a matter of fact, laughter commonly occurred throughout the conversations, for which we identified three primary reasons. First, participants laughed when reminiscing about pleasant times or describing experiences with joy and pride. For example, P10 chuckled with delight when talking about her son being admitted to his dream university. Second, they laughed when triggered to recall funny events from their past. P6, for example, chuckled while telling the story of an early mobile phone user in her neighborhood being mocked for seemingly talking to a "brick". Third, they also found amusement in their interactions with the

agents when the agents posed naive questions (e.g., P8 burst into laughter when the child agent asked if she had ever hung stuff from a ceiling fan) or made silly mistakes in recognizing people in the photos (e.g., P3).

Participants also expressed a willingness to use the system in their daily lives (Mean = 4.3, Median = 4.0). Interviews indicated three main reasons for the positive reception. First, the system assisted participants in recalling past experiences through storytelling (P3, P4, P6, P7). During the interview, P5 shared an anecdote about her friend that echoed this statement. During the pandemic, this friend, approximately 60 years old, created a slideshow of family photos for his father, who was nearly 90. Day after day, he would play this slideshow for his father and guide his father in sharing stories about each photo. This ritual became their way of passing time and maintaining connection during the lockdown periods, reflecting the power of reminiscence through storytelling.

Second, participants appreciated having a “*person*” to converse with while reminiscing (P1, P5, P9), which “enhanced their willingness to share stories” (P11). This was particularly valued by older adults living alone. As P9 explained, “*I like to have my TV on when I'm alone at home – as far as I know, many people like me do this – having some sounds at home makes me feel accompanied.*”

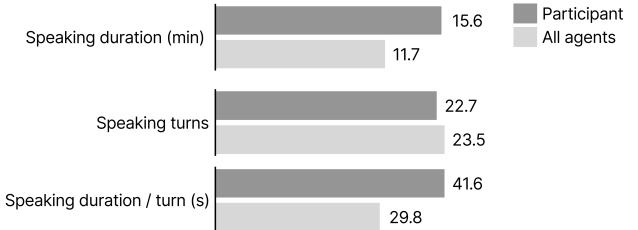
Third, the system allowed for independent reminiscence without the presence of others. P3 elaborated, “*I didn't need to consider how the agents think or respond, without purposes such as impressing others or educating younger generations. It is a fully self-reflective process.*”

## 5.2 User-Agent Conversation Dynamics

To investigate user-agent conversation dynamics, we first examined turn-taking behavior to assess the balance of conversation and how agents and participants respond to each other. Next, we analyzed participants' response content to explore how the conversations influenced reminiscence. Lastly, we investigated participants' perceptions and attitudes toward the conversations.

**5.2.1 Turn Taking Behavior.** During the prototype trial across the two themed sessions, each participant engaged in conversations for an average of 15.6 minutes, comprising a mean of 22.7 turns, with each turn lasting an average of 41.6 seconds; meanwhile, the four agents spoke for a total of 11.7 minutes (averaged across participants), taking 23.5 turns, each turn lasting around 29.8 seconds (Figure 7). These results suggest a relatively balanced dynamic of interaction between participants and agents, indicating that the system has successfully fostered natural and engaging conversations.

Analysis of participant responses revealed three response patterns: *reactive response*, *active response*, and *proactive initiation*, with examples listed in Figure 8. *Reactive response* refers to participants directly answering the agents' questions, sometimes including brief but minimal social engagement with the agents, varying from short to more extended answers. *Active response* involves participants socially engaging with the agents, sometimes asking about the agents' personal backgrounds or experiences. *Proactive initiation* occurs when participants initiate new topics, which could involve knowledge-based discussions or social interactions with the agents. Overall, participants spent more time responding to the agents,



**Figure 7: Comparison of turn-taking behavior between participants and *Responding Agents*, regarding speaking duration, number of speaking turns, and average speaking duration per turn.**

compared to actively engaging with the agents or starting new topics, as the distribution shown in Figure 8, which shows the number of people exhibiting different patterns.

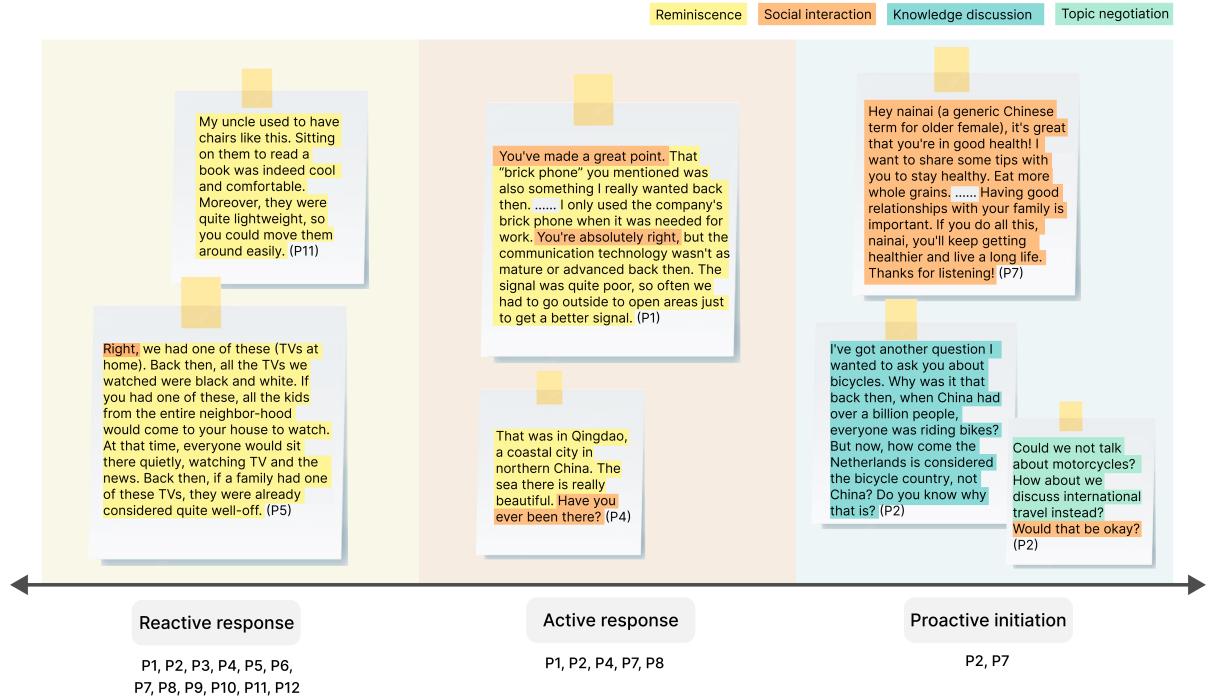
Eye-tracking data revealed that users were generally engaged, as most of their eye gaze focused on the regions of interest (the artifact and the two agents)(Figure 9). During interactions, participants looked at the displayed artifact significantly longer than at the two agents combined ( $t_{11} = 5.20, p < 0.001$ ).

**5.2.2 Distribution of User Response Content.** We analyzed the participants' response contents and identified seven categories: *Autobiographical memory*, *Public memory*, *Factual memory*, *Comparison with present*, *Social interaction*, *Knowledge discussion*, and *Topic negotiation*. The explanation and examples for each category are listed in Table 2.

The distribution of participants' response contents is illustrated in Figure 10, arranged in descending order based on the word (Chinese character) count. The large word counts and high occurrence for *Autobiographical* and *Public memory* categories indicate that reminiscence was effective. Moreover, there is a notable portion of *Comparison with present* responses, suggesting participants reflected on current living conditions, mainly positively according to post-interview. Additionally, participants constantly engaged in social interactions with the agents, albeit with fewer words, demonstrating their willingness to converse with the AI agents.

**5.2.3 Conversational Experience and Functions.** Overall, participants reported that interacting with the system was fluent and coherent (P5, P10, P12). They were particularly impressed by the agents' capacity to understand and engage in smooth and natural conversations (P1, P5, P8, P9, P10, P11). Participants rated they liked conversing with the four agents (Mean = 4.5, Median = 4.0). P8 even remarked, “*I initially thought you were controlling the agents' responses behind the scenes.*”

Interviews indicated that all participants endorsed the conversational approach of the system, with the following three main reasons. First, this approach provides a conversation partner that guides the storytelling process. The conversational format facilitated the initiation of storytelling, providing participants a starting point, “*making the whole process easier to start*” (P8). Moreover, the interactive approach encouraged participants to “*share more contents, which is a good form of language exercise*” (P1) and “*can stimulate the brain to be more active*” (P3).



**Figure 8: Participant response patterns.** Three types of responses were identified: (1) **Reactive response**: participants directly answer agents' questions; (2) **Active response**: participants engage socially with agents, sometimes inquiring about their experiences; (3) **Proactive initiation**: participants introduce new topics or initiate knowledge-based discussions. The figure illustrates examples of each response type.



**Figure 9: (A) Heatmap of all participants' gaze distribution. (B) Percentage of gaze distribution over ROIs.**

Second, participants appreciated how the agents would first provide responses, usually positive, to what they said before continuing the conversation. *"This made me feel very delighted."* (P6) They also reported an inclination to engage with agents that provided more positive feedback during conversations, as exemplified by P1's stated preference for the boy agent *"the boy agent response with positive feedback and spoke in a manner that was more flattering"*. Yet, participant (P2) pointed out that "It's not always good for responses to be positive. I would prefer them to be more targeted."

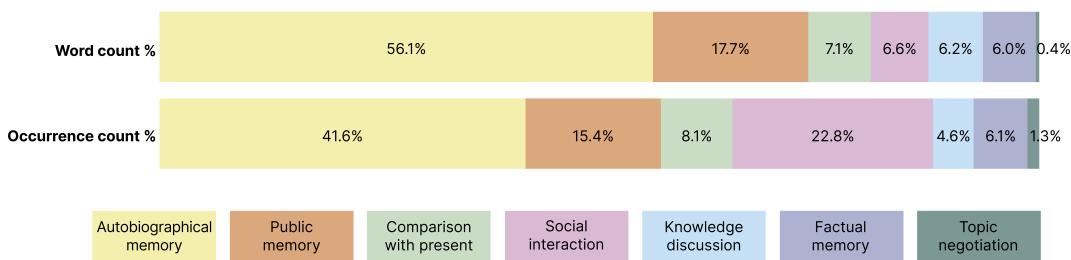
Third, conversing with agents offered participants new aspects in their reminiscence processes. The agents' descriptions, associative

thinking, and inquiries about artifacts triggered *"more memories and recalled more details"* (P12) than they might have when simply narrating stories about the artifacts directly. P9 noted, *"The agents' detailed descriptions and questions helped me remember things I hadn't initially recalled. For example, when it came to TV, I might have only remembered how bulky they were at first. But when he (the agent) asked about the antenna, it triggered memories of how we used to adjust it to get a better signal."*

Last, conversing with agents offers additional benefits compared to interacting with a real person. While human conversation partners are not always available, agents can be accessible at any time.

**Table 2: Response content categories and examples.**

No.	Content category	Explanation	Examples
1	<i>Autobiographical memory</i>	Recollections of one's own life experiences and personal events.	"That winter wasn't particularly cold, and it didn't snow. So, during the festive atmosphere of the New Year, our whole family went out to enjoy ourselves in the park. (P3)"
2	<i>Public memory</i>	Recollections of previous lifestyles.	"Back in those days, if one family had a television, all the neighbors would gather to watch together. Ever since we got TV, our cultural life has become much more enriched. (P12)"
3	<i>Factual memory</i>	Descriptions of the displayed artifact, such as who is the person in the photo or the details of the object.	"The one being held is my granddaughter, and the person holding her is my elder brother's daughter. (P6)" or "That thing on top is the flash bulb. We used it when taking photos at night or in low-light conditions. (P9)"
4	<i>Comparison with present</i>	Comparisons between what things were like in the past and how they are now.	"Nowadays, we don't do that (sewing clothes) anymore. There's no need for it now. These days, we just go to the shopping mall to buy clothes. (P5)"
5	<i>Social interaction</i>	Social responses to the agents, such as praising agents or asking about the agents' personal experiences.	"Haha, what you said is really interesting! (P4)" or "Have you also been there? (P4)"
6	<i>Knowledge discussion</i>	Exchanging opinions or asking questions with the agents.	"Each type of transportation has its own unique features. None can completely replace the others. Everyone has different travel habits and needs. (P2)"
7	<i>Topic negotiation</i>	Negotiating with the agents about the topic under discussion.	"How about we discuss international travel instead? (P2)"

**Figure 10: The distribution of participants' response contents.**

Furthermore, users can express themselves freely when talking to an agent, without concerns about judgment or reactions, which is often not the case when talking to a real person. As P2 noted, “*Even if I’m angry, I can just yell at it without worrying consequences.*” Consequently, this process may facilitate self-reflection, allowing people to be more candid about their inner world, and potentially serve as a means of emotional regulation.

However, a common concern among participants was the lengthy response time of the agents. On average, participants waited 31 seconds ( $SD = 6.8$ ) between finishing their speech and hearing the

agent’s response. P6 commented, “*While I can understand the system needs time to process, the waiting time feels awkward.*” P1 also stated, “*the long wait kind of killed my enthusiasm to continue.*” After the study, we conducted additional analysis by simulating conversations and examined the factors that contributed to participants’ waiting time. Results showed that the delays were primarily caused by participants’ speech transcription and *Responding Agents’* responses generation, particularly for long participant utterances. Speech recognition time varied with the length of the utterances, ranging from 5 to 15s. The generation time for *Responding Agent’s*

responses increased from 4 to 20s as conversation history grew, which increased the context content that was fed to LLM to generate responses. The *Artifact Switching Agent* decisions required about 3s, while converting *Responding Agents*' responses to speech took about 2.5s. Remaining time was due to program execution and network transmission. Moreover, participants wished for agents to intelligently determine who the user intended to converse with. Although some participants found the current eye tracking method “convenient” (P9), others considered it not natural enough. For example, P4 noted the discomfort of feeling obligated to maintain eye contact with an agent for the sake of politeness while responding to their question, which consequently led to a prolonged conversation with that particular agent.

Some participants also noted a lack of depth in the dialogues despite the overall fluency of the conversations, probably due to the issues with topic management. P4 observed, “*Just when I’m really getting into it, the topic changes, and we end up barely scratching the surface of anything.*” P11 also noted that “*sometimes the agents would ask questions about trivial details, such as what is the element on a displayed object, which is a little bit meaningless.*”

### 5.3 Perspectives on Agent Design

Overall, participants expressed a positive perception of the agents, rating them favorably in terms of overall likability, conversation naturalness, conversation fluency, conversation effectiveness, and trustworthiness (Figure 11). Wilcoxon signed-rank tests showed that ratings between older and younger agents did not significantly differ for all the five dependent variables. Comparative analyses of word counts across the seven categories of response content, conducted using Wilcoxon signed-rank tests, also demonstrated no significant differences in the distribution of response content between the older and younger agents (Figure 12). Eye-tracking data also showed no significant difference in the percentage of time users spent looking at older agents versus younger agents ( $W = 33, p = 0.677$ ).

Although the ratings and the distribution of response content between the older agent and the younger agent did not statistically differ, participants stated that they treated the two types of agents differently. As one participant (P2) noted, the choice of agents “*depends on my mood at the time and the topics I want to discuss. ... When chatting with Xiaoming (the boy agent), I was more likely to explain about how things were in the past. But with the older agent, knowing that he has also experienced those times, I would directly discuss the situation at that time with him.*”

Wilcoxon signed-rank tests revealed significant differences in participants’ conversation rounds (older agent = 10.4, younger agent = 13.0,  $W = 7, p = 0.035$ ) and speaking duration (older agent = 6.2 min, younger agent = 9.6 min,  $W = 8, p = 0.012$ ) between interactions with older and younger agents, demonstrating a tendency for participants to engage more frequently with the younger agent. Participants’ speaking duration per round did not significantly differ between the two types of agents (older agent = 45.4s, younger agent = 46.2s,  $W = 35, p = 0.791$ ). Post-survey interviews revealed that the primary reason for this preference was that the younger agents were perceived more **emotionally engaging**. Participants reported finding the younger agents more appealing, describing

them as “*adorable*” (P4, P11, P12), which fostered “*a sense of closeness*” (P12). They were perceived as having “*a brighter future*” (P6), with participants being “*attracted to their vibrant energy*” (P3). Moreover, the younger agent’s responses were perceived as “*emotionally richer, making me want to keep talking to it*” (P1). Interestingly, these impressions were not solely derived from our prototype’s avatar design but also influenced by participants’ real-life experiences. As one participant (P10) expressed, “*It’s not only these child agents but also children in general. I’ve always had a preference for conversing with children in my daily life. It makes me feel younger.*” P5 noted that “*sharing with the child agents made me feel my era and my own experiences were more meaningful.*”

On the other hand, the older agent was perceived as **more relatable in content**, sharing both background knowledge and spiritual resonance. Some participants preferred conversing with the older agent due to the perception of “*shared more background knowledge (than the younger agent)*” (P9) and finding it “*easier to communicate*” (P1). As P9 noted, “*When talking to the child agent, it responded shallowly, making me wonder whether it truly understands. Feels like there was a generational gap.*” Furthermore, participants stated that interacting with the older agent fostered “*a stronger sense of mutual communication, characterized by sharing, empathy, and reciprocal understanding*” (P4). Participants felt they could “*share and discuss life insights with the older agent, whereas discussing such topics with a child feels weird as they might not understand*” (P3).

**5.3.1 Expectations for Agent Design.** Regarding agent identity, interviews revealed that participants were generally hesitant about agents mirroring individuals from their personal lives. First, participants believed the psychological projection involved would influence the content and emotions of their reminiscences. For instance, due to certain contents of reminiscences that are not suitable for sharing with acquaintances, such as unpleasant experiences, they would not share these with the agent mimicking that familiar person. As P10 explained, “*There’s always a bit of conflict between people, which could potentially bring back some unpleasant memories.*” This could cause them to be cautious when speaking, worrying that the agent might not want to listen (P1, P3). Second, there were concerns about the agent’s prior knowledge being inadequate, especially regarding the specific details of shared experiences, which hindered its ability to maintain a high level of authenticity during the reminiscence process. As P2 stated, “*If agents just look like someone but lack the shared experiences I have with that person, they might come off as insincere and out of place.*” Conversely, participants were receptive to agents embodying celebrities or characters related to the discussion topic. They believed interacting with celebrities or well-known characters could make reminiscence more interesting and engaging, as stated by P10 that “*It would be really interesting to chat about my sports experiences with Olympic winners.*”

In terms of the appearance of the agents, participants found the agents to be dull, lacking in vividness, and suggested designing more vivid appearances. The current cartoonish representation of the agents was perceived as not engaging enough. As P9 elaborated, “*The design of these agents creates a sense of psychologically distant, making me worried that they were reluctant to talk with me.*” Furthermore, participants suggested that more detailed design elements should be added to the agents to make them vivid and realistic.

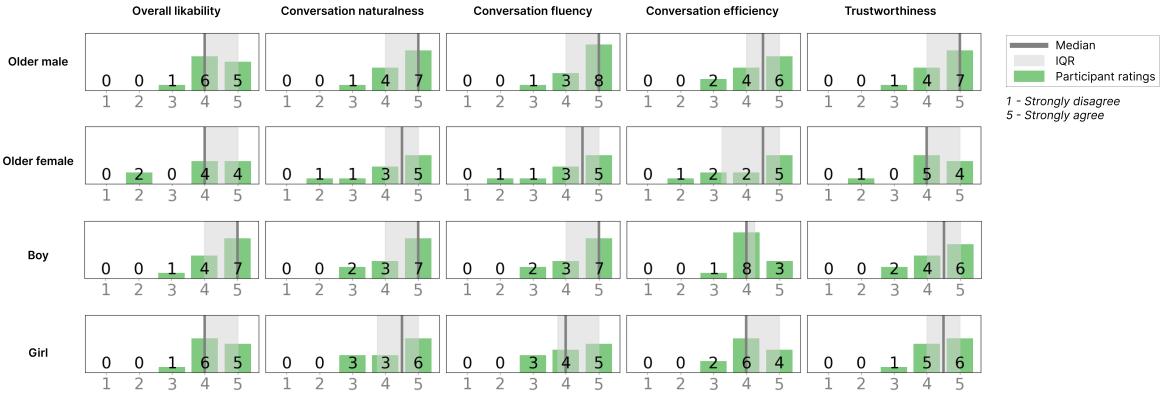


Figure 11: Participants' ratings for the four Responding Agents.

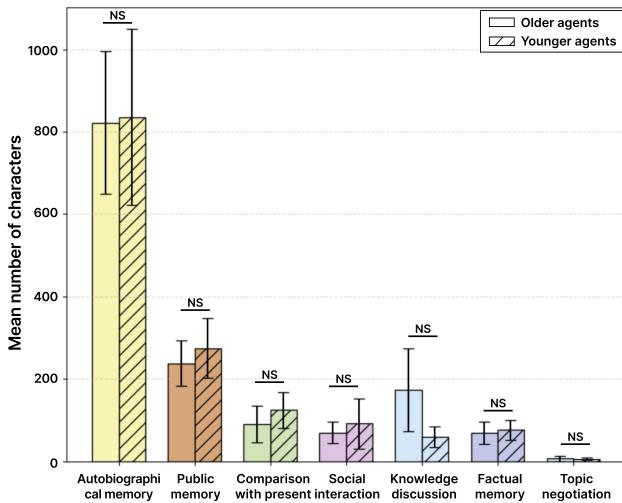


Figure 12: Mean number of Chinese characters per response content category with older and younger agents. Error bars represent standard error; NS: not significant.

From the verbal behavior perspective, they proposed varying the tone according to the emotional content (P11). From the nonverbal behavior perspective, participants suggested incorporating body movements (P4), eye contact feedback (P9), and facial expressions (P5), emphasizing nonverbal behaviors' importance in delivering human-like feelings.

#### 5.4 Perspectives on Artifacts

In general, participants reported that both personal photos and 3D generic nostalgic objects displayed during conversations were appropriate, facilitated the recall of past experiences, enhanced their engagement in conversations, and increased their willingness to use the system (Figure 13). Wilcoxon signed-rank tests showed that ratings between 3D generic nostalgic objects and personal photos did not significantly differ for all the four dependent variables. These artifacts effectively guided older adults in reminiscence activities.

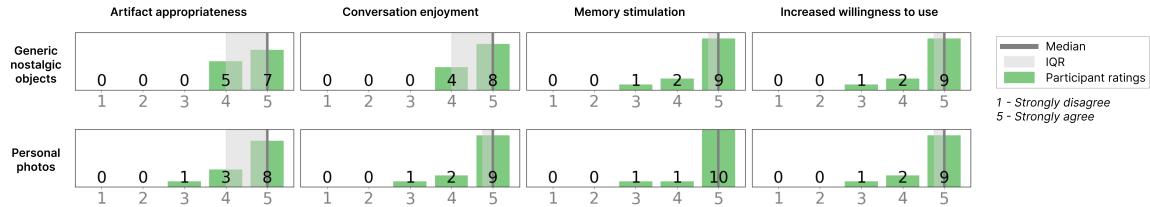
As one participant noted, they were all “*precious memory treasures*” (P6), while another mentioned that the objects “*touched my heart*” (P3).

We did not analyze the differences in participants' total speaking duration or conversation rounds between personal photos and 3D generic nostalgic objects, as these were influenced more by the decisions of the *Artifact Switching Agent* than by the participants themselves. Participants' speaking duration per round did not significantly differ between the two types of artifacts (personal photo = 43.4s, 3D generic nostalgic object = 48.7s,  $W = 25, p = 0.301$ ).

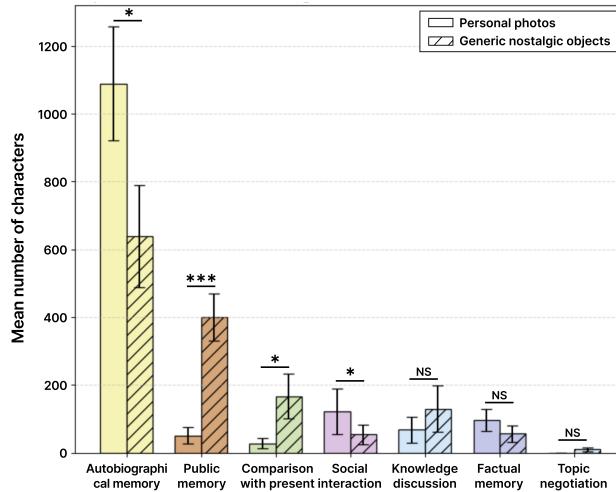
Analysis of the response content distribution between the generic nostalgic objects and personal photos showed that the two types of artifacts elicited different forms of reminiscence (Figure 14). We conducted a comparative analysis of word counts across the seven categories of response content, using the Wilcoxon signed-rank test, contrasting personal photos with generic nostalgic objects. The results showed that personal photos, compared to generic nostalgic objects, elicited more *Autobiographical memories* ( $W = 9, p = 0.016$ ), and less *Public memories* ( $W = 0, p < .001$ ), less *Comparisons with present* ( $W = 11, p = 0.050$ ), and more *Social interactions* ( $W = 10, p = 0.021$ ). As P10 stated, “*photos evoke memories of people and events, while the old objects remind me of the era itself*.”

First, personal photos predominantly evoked *Autobiographical memories*, including individual experiences, significant life events, the evolution of relationships, etc. Personal photos could “*instantly trigger more memories, providing a wealth of stories to tell*.” (P2) Moreover, participants enjoyed the process of reminiscing about the past through old photos, stating that the system helped to better utilize the value of these old photos. As P12 expressed, “*Most of the photos we typically browse are from recent years, but spend less time looking through the old ones*.”

In contrast, generic nostalgic objects tended to elicit shared memories of the era, representing common experiences across a generation. Participants expressed greater surprise when seeing these objects, as it is “*rare to come across these objects in daily life*,” (P5) and “*such objects are seldom captured in photos*.” (P11). This statement was echoed by many participants during our recruitment process saying that photos containing the objects we specified were rather difficult to find. Additionally, seeing these objects made participants



**Figure 13: Participants' ratings for generic nostalgic objects and personal photos.**



**Figure 14: Mean number of Chinese characters per response content category for personal photos and generic nostalgic objects. Error bars represent standard error; \* p < 0.05, \*\*\* p < 0.001, NS: not significant.**

“feel very close and familiar” (P5), allowing them to relive the era they represented, “triggering lots of memories” (P1, P2). As P6 expressed, “They practically transported me back to that era, making me feel very excited.” Furthermore, the generic nostalgic objects often prompted Comparisons with present, leading to reflections on societal changes. As P9 expressed, “When I recall these details from the past, it really makes me nostalgic. Thinking about all we’ve been through, and how we’ve arrived at this high-tech life now, I feel a real sense of joy and pride in how far we’ve come.”

We also found that the effectiveness of generic nostalgic objects depends on the participants’ familiarity with the item, as well as the richness of their experiences related to it. Participants were able to share more extensively about objects they had previously used and had stories associated with. These narratives often included details such as when they first encountered the object, how it was used, who they used it with, and interesting anecdotes related to it. Conversely, when participants encountered objects they were unfamiliar with, they found it relatively challenging to elaborate on them, with P2 noting that there was “nothing much to say.”

In addition, the 3D representation further enhanced the impact and stimulating effect of the generic nostalgic objects. Participants all found the models to be very realistic, “as if the object is truly present in front of you.”(P6) Most of the participants inquired whether they could touch and play with the model during the study. Throughout the conversations, they frequently explored the displayed model by viewing it from various angles. As P11 summarized, “The 3D format has a really powerful visual impact. It leaves a deep impression.” In the post-study survey, participants all rated *strongly agree/agree* with the statement “not feeling sick when using the system” (Mean = 4.8, Median = 5.0), indicating no cybersickness.

## 6 Discussion

With the rapid advancement of LLMs, LLM-powered agents show great potential in engaging older adults in conversations flexibly around various memory cues to assist in reminiscence. However, there’s a significant research gap in how to design such systems and understand their impact on reminiscence behavior. Specifically, how the conversation dynamics should be designed, and how agent identities and various artifacts as memory cues would affect the reminiscence behavior remain underexplored. To address this research gap, we presented ReminiBuddy, a multi-agent reminiscence system, featuring agents with older and younger identities, as well as generic nostalgic objects and personal photos as memory cues. We used it as a technology probe to explore how it might facilitate reminiscence in older adults. We first examined the effectiveness of the conversational approach in supporting reminiscence among older adults. We then investigated how different agent roles and types of artifacts influence the reminiscence process. Overall, we found that conversational interactions between the agent and older adults effectively facilitated reminiscence. Furthermore, the diversity of agents and memory cues played complementary roles in the reminiscence process, enabling older adults to reminisce from various perspectives and enriching the content of their recollections. In this section, we discuss our findings and key design implications to guide future practices towards effectively supporting reminiscence in older adults using LLM-powered conversational agents.

### 6.1 Future Design Considerations and Directions

Our study offers the following design considerations to inform the future development of similar applications, with a focus on

agent identity and design, the use of artifacts as memory cues, and fostering more natural and flexible conversational dynamics.

**6.1.1 Exploring the full spectrum of agent design in enriching the reminiscence experience.** Our study involved four agents with distinct identities, varying in age and gender, achieved through differences in appearance and prompt design. We found that older adults' interaction behaviors, perceptions, and preferences varied with agent identities, aligning with previous research showing that metaphorical descriptions help older adults conceptualize voice assistants (VAs) and establish interaction norms [11, 17]. Specifically, interactions with younger agents prompted more conversational turns and longer engagement than with older agents. Younger agents were described as energetic and emotionally engaging, consistent with findings that emotional support motivates older adults to engage in reminiscence [49]. Participants often likened younger agents to children, with comments like, "Sharing knowledge with children who don't know about it is meaningful," reflecting real-life reminiscence motivations [15]. In contrast, older agents were seen as more relatable and suitable for reflective conversations on life experiences and relationships. These results suggest that distinct agent roles can enhance reminiscence by addressing varied needs: younger agents may foster emotionally uplifting interactions, while older agents facilitate deeper, self-reflective dialogues, such as reframing negative experiences into positive insights [31]. While we could not tell the differences in interaction behavior between agent identities were due to variations in visual appearance, voice characteristics or prompt design, we prioritized consistency in the agents' attributes to its identity. Controlled studies with comprehensive metrics, such as user experience, emotional responses, and reminiscence depth, can further clarify how agent identities and their specific attributes affect older adults, guiding the design of personalized reminiscence technologies and better addressing the nuanced needs of older adults during reminiscence activities. We did not explore the effect of the gender of agents, although we observed a slight preference for the male agents in our experiment, due to the small sample size and the slight gender imbalance among participants. Future research could also consider factors beyond age and gender, such as cultural background and personality traits, to further diversify agent identities. With the rapid advancements in LLMs capable of role-playing (e.g., [33, 48]), this becomes increasingly feasible.

Moreover, unlike traditional VAs, which are often perceived as fictional characters or acquaintances [11], our agents were mostly viewed as distinct individuals. For instance, one participant engaged with an older agent asked about its life and expressing concern for its well-being. However, this mental model was not consistent, especially regarding the older agent. Some saw the agents as peers, while others associated them with their own grandmother. These diverse, anthropomorphic interpretations highlight the complex mental models that agents with human-like qualities evoke. Furthermore, the metaphorical description of agents can influence participants' impressions and may lead to negative effects on conversation topics and feelings of relaxation. Some participants felt constrained when interacting with older agents perceived as peer strangers, comparing these interactions to real-life conversations with adults. This sometimes discouraged sharing personal stories due to privacy

concerns. For the same reason, some participants found agents modeled after family or friends undesirable, as they felt more restricted in their choice of topics, worrying about whether the agents would "want" to listen. Thus, when designing the agents assisting older adults in reminiscence, we recommend that the agent identity also need to be carefully designed. Future work should explore the boundaries between viewing agents as strangers, familiar figures, or surrogate family members, and how these perceptions affect the depth and quality of reminiscence. This exploration could offer key insights into how to fine-tune agent personas to align with different reminiscence goals, from light social engagement to deep emotional processing.

**6.1.2 Investigating and implementing diversified and relevant artifacts as memory cues.** Our study explored the use of both personal photos and generic nostalgic objects as memory cues, uncovering distinct patterns in their effectiveness. Personal photos predominantly elicited autobiographical memories (e.g., "When I was..."), while generic nostalgic objects were more effective in evoking era-related memories, such as public recollections ("Everyone back then...") and comparative reflections ("Back then... but now..."). These findings align with prior research demonstrating the efficacy of personal and generic photos in supporting reminiscence [2, 6]. However, our work builds on these insights by demonstrating that generic objects uniquely support the recall of public and comparative memories, often linked to broader historical or cultural contexts, while introducing new reminiscence scenarios where the conversational partner is an agent rather than a human. This highlights their potential to enrich the reminiscence process beyond personal storytelling. Generic artifacts also serve as an effective alternative in scenarios where personal photos are unavailable, enabling memory recall and storytelling [49]. Moreover, engaging with generic objects may involve greater cognitive effort to retrieve forgotten details, providing an additional layer of cognitive stimulation—a recognized benefit of reminiscence activities [8, 13]. While prior technologies have predominantly emphasized personal photos as memory prompts, particularly for storytelling with family members or peers [7, 28], our findings suggest that generic objects play a complementary role. Thus, the choice of artifacts should align with the specific goals of reminiscence—whether the focus is personal storytelling, cognitive training [39], or engaging activities such as games [1].

However, our study also highlighted a key limitation: participants emphasized that effective reminiscence materials must resonate with their lived experiences. Objects disconnected from their personal history or cultural context often failed to inspire meaningful engagement. While our study used a limited set of generic artifacts, future research could expand this scope by developing larger, more adaptable collections tailored to individual preferences. Advancements in generative AI present exciting opportunities for enhancing memory prompts. Recent studies by Jin and Cai et al. explored generating images based on older adults' conversations to support reminiscence and storytelling [22]. Although challenges remain in aligning generated content with participants' memories, future systems could dynamically adjust nostalgic objects—modifying colors, adding elements, or generating personalized scenarios—to better match users' experiences. Such advancements could create

more flexible and accessible memory prompts, ultimately enhancing the effectiveness and personalization of reminiscence technologies for older adults.

**6.1.3 Advancing toward more natural and flexible conversational dynamics.** Our study on designing conversational experiences for older adults with two agents yielded valuable insights, with participants offering generally positive feedback. The interactions revealed important strengths in the system, particularly in terms of engagement and relevance to their needs. However, our analysis identified areas for improvement, particularly in the dynamics of interactions. These findings inform potential design implications and future directions:

**(1) Adapting to older adults' interaction patterns.** Our study identified three interaction patterns among participants: reactive, active, and proactive initiation. Most participants exhibited reactive behaviors, responding primarily to the agents' prompts. Some participants actively engaged by offering compliments, asking follow-up questions, or introducing new topics, such as personal concerns or opinions about the agents. Proactive participants occasionally requested shifts to more interesting topics. However, the agents' rigid artifact-centered prompts limited these opportunities, diminishing the potential for dynamic and personalized exchanges. This aligns with prior research highlighting non-cooperative user behavior, where users deviate from predefined paths when dissatisfied with the conversation topics [27]. To address this, conversational designs should accommodate varied interaction styles by allowing prompts to adapt dynamically. For example, when participants deviate from artifact-centered discussions to share personal reflections or request new topics, agents should recognize and respond to these cues rather than redirecting back to predefined structures. Implementing mixed-initiative dialogue systems [16], where both users and systems flexibly lead the conversation, could enhance personalization and create more meaningful exchanges. Such flexibility would align with older adults' conversational preferences and improve engagement.

**(2) Enhancing artifact and topic switching.** Currently, artifact-switching prompts are uniform, with predefined conversation lengths of approximately five turns per artifact. While this ensures exposure to multiple artifacts, it limits conversational depth and adaptability. Participants often expressed frustration when artifacts they found particularly interesting were switched too quickly, whereas less engaging artifacts received disproportionate attention. To improve this, artifact-switching mechanisms should leverage multimodal cues such as engagement duration, voice tone, or sentiment analysis to gauge participant interest more effectively. For example, if a participant expresses enthusiasm or shares detailed stories about an artifact, the system should extend the interaction. Conversely, when participants show disinterest or provide minimal input, the system could shorten the discussion and transition to a more engaging artifact. This dynamic approach would keep conversations relevant and engaging, catering to participants' preferences.

**(3) Refining the turn-taking mechanism.** The current turn-taking mechanism, based on eye-tracking, introduced a level of natural interaction but also revealed limitations. Some participants found prolonged eye contact uncomfortable and made eye contact

out of politeness rather than an expectation of a response. This occasionally resulted in unintended conversational turns, disrupting the flow and reducing the depth of engagement with a single agent. Future turn-taking mechanisms should incorporate additional cues such as verbal acknowledgments, subtle gestures, or conversational pauses to facilitate more intuitive interactions [34]. This approach would reduce misunderstandings and allow participants to engage at their own pace, free from the pressure of rigid design constraints. Moreover, turn-taking should not be limited to participants and agents. Adding interactions between the agents themselves could address moments when conversations become less engaging or participants struggle to provide relevant input. Conversations between agents could serve as an additional memory cue for older adults, stimulating reminiscence and creating a more realistic social interaction scenario. This approach could further satisfy older adults' social needs, making interactions with the system feel more authentic and dynamic.

## 6.2 Limitations

Our work has limitations that could potentially motivate future research directions. First, due to the exploratory nature of this study, we couldn't identify the specific effects of certain variables, such as differences between 2D and 3D formats, the varying familiarity participants had with generic objects versus personal photos, and generic object and photo content. Specifically, the 3D effect seemed to be engaging, as it appeared to stimulate participants' interest and desire to interact with the generic nostalgic object. This raises the intriguing question of whether similar engaging effects could be achieved by rendering personal photos in 3D. Such a development could potentially enrich the reminiscence process, as the immersive quality of 3D visuals might trigger more vivid emotional and cognitive responses when interacting with personal, meaningful memories [28]. Exploring this possibility could reveal subtle nuances in how 3D visuals influence the recall and emotional engagement with personal photos, thereby offering a deeper understanding of their potential impact on reminiscence. A more controlled study is needed in order to explore these factors' impact on the reminiscence process and effectiveness. Second, our participants came to the lab to try the system for relatively short periods. Open questions remain about how they might use such a system over longer terms and in real-world settings. For example, it's unclear how people would feel about having AI agents in their homes, or how the system might integrate into daily routines. Future work should address the generalizability of these aspects. Third, cultural differences may affect how older adults perceive and interact with the system. The role of reminiscence, attitudes towards technology, and inter-generational communication styles, the impact of different artifacts and photo content on user engagement can vary across cultures [29]. Future research needs to go further to include participants from diverse cultural contexts. Last, participants in our study were all relatively well-educated, open-minded, and possessed a certain level of proficiency towards new technologies. How less-educated older adults or those more conservative towards new technologies might perceive this system remains an open question. Incorporating these demographics into future studies is crucial to inform the development of such systems for broader use.

## 7 Conclusion

The advancement of LLMs has opened new possibilities for AI-driven reminiscence support, enabling dynamic interactions between AI agents and older adults using diverse memory cues. However, the optimal design of such systems and their effects on reminiscence practices remain unexplored, specifically regarding the design of conversation dynamics, agent identities, and artifacts as memory cues.

To address this gap, we present ReminiBuddy, a novel LLM-powered multi-agent system enabling older adults to interact with two distinct agents—one older, one younger—using personal photos and 3D models of generic nostalgic objects as memory prompts. Conversations with peers often spark a sense of shared experience [7], while storytelling with younger generations has demonstrated numerous benefits for reminiscence [28, 30, 42]. As for memory cues, while personal photos are the most frequent cues, research has shown that generic artifacts can sometimes bring additional benefits [2]. Therefore, we incorporated these diverse elements into our system to inspire future design of such systems. Results showed that participants found the conversational process enjoyable and beneficial for reminiscence. The younger agent was perceived as more emotionally engaging, while the older agent was more relatable in content. Personal photos triggered autobiographical memories, whereas generic nostalgic objects evoked shared era-related memories, collectively offering a multifaceted reminiscence experience.

Based on these results, we propose design implications to inform the future development of conversational reminiscence systems. We also discuss future research directions for building such systems, aiming to enhance their effectiveness and user experience for older adults.

## Acknowledgments

This work is partially supported by CCF-Lenovo Blue Ocean Research Fund (No.: CCF-Lenovo OF 202302), the Guangzhou-HKUST(GZ) Joint Funding Project (No. 2024A03J0617), Guangzhou Higher Education Teaching Quality and Teaching Reform Project (No. 2024YBJG070), Education Bureau of Guangzhou Municipality Funding Project (No. 2024312152), Guangdong Provincial Key Lab of Integrated Communication, Sensing and Computation for Ubiquitous Internet of Things (No. 2023B1212010007), the Project of DEGP (No.2023KCXTD042), and the Guangzhou Science and Technology Program City-University Joint Funding Project (No. 2023A03J0001).

## References

- [1] Alaa Abd-alrazaq, Dari Alhuwail, Eiman Al-Jafar, Arfan Ahmed, Farag Shuweihdi, Shuja Mohd Reagu, and Mowafa Househ. 2022. The Effectiveness of Serious Games in Improving Memory Among Older Adults With Cognitive Impairment: Systematic Review and Meta-analysis. *JMIR Serious Games* 10, 3 (Aug. 2022), e35202. <https://doi.org/10.2196/35202>
- [2] Arlene J. Astell, Maggie P. Ellis, Norman Alm, Richard Dye, and Gary Gowans. 2010. Stimulating people with dementia to reminisce using personal and generic photographs. *International Journal of Computers in Healthcare* 1, 2 (2010), 177. <https://doi.org/10.1504/IJCIH.2010.037461>
- [3] Benett Axtell and Cosmin Munteanu. 2018. Frame of Mind: Using Storytelling for Speech-Based Clustering of Family Pictures. In *Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion*. ACM, Tokyo Japan, 1–2. <https://doi.org/10.1145/3180308.3180330>
- [4] Benett Axtell and Cosmin Munteanu. 2019. Back to Real Pictures: A Cross-generational Understanding of Users' Mental Models of Photo Cloud Storage. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (Sept. 2019), 1–24. <https://doi.org/10.1145/3351232>
- [5] Benett Axtell and Cosmin Munteanu. 2019. PhotoFlow in Action: Picture-Mediated Reminiscence Supporting Family Socio-Connectivity. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland UK, 1–4. <https://doi.org/10.1145/3290607.3313272>
- [6] Benett Axtell, Raheleh Saryazdi, and Cosmin Munteanu. 2022. Design is Worth a Thousand Words: The Effect of Digital Interaction Design on Picture-Prompted Reminiscence. In *CHI Conference on Human Factors in Computing Systems*. ACM, New Orleans LA USA, 1–12. <https://doi.org/10.1145/3491102.3517692>
- [7] Steven Baker, Ryan M. Kelly, Jenny Waycott, Romina Carrasco, Roger Bell, Zaher Joukhadar, Thuong Hoang, Elizabeth Ozanne, and Frank Vetere. 2021. School's Back: Scaffolding Reminiscence in Social Virtual Reality with Older Adults. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW3 (Jan. 2021), 1–25. <https://doi.org/10.1145/3434176>
- [8] Ernst Bohlmeijer, Marte Roemer, Pim Cuijpers, and Filip Smit. 2007. The effects of reminiscence on psychological well-being in older adults: A meta-analysis. *Aging & Mental Health* 11, 3 (May 2007), 291–300. <https://doi.org/10.1080/13607860600963547>
- [9] Mariona Carós, Maite Garolera, Petia Radeva, and Xavier Giro-i Nieto. 2020. Automatic Reminiscence Therapy for Dementia. In *Proceedings of the 2020 International Conference on Multimedia Retrieval*. ACM, Dublin Ireland, 383–387. <https://doi.org/10.1145/3372278.3391927>
- [10] Qian Chen, Mengzhe Chen, Bo Li, and Wen Wang. 2020. Controllable Time-Delay Transformer for Real-Time Punctuation Prediction and Disfluency Detection. In *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 8069–8073. <https://doi.org/10.1109/ICASSP40776.2020.9053159>
- [11] Jessie Chin, Smit Desai, Sheny (Cheng-Hsuan) Lin, and Shannon Mejia. 2024. Like My Aun Dorothy: Effects of Conversational Styles on Perceptions, Acceptance and Metaphorical Descriptions of Voice Assistants during Later Adulthood. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW1 (April 2024), 1–21. <https://doi.org/10.1145/3637365>
- [12] Martin A. Conway. 2009. Episodic memories. *Neuropsychologia* 47, 11 (Sept. 2009), 2305–2313. <https://doi.org/10.1016/j.neuropsychologia.2009.02.003>
- [13] Maria Cotelli, Rosa Manenti, and Orazio Zanetti. 2012. Reminiscence therapy in dementia: A review. *Maturitas* 72, 3 (July 2012), 203–205. <https://doi.org/10.1016/j.maturitas.2012.04.008>
- [14] Sara J. Czaja, Neil Charness, Arthur D. Fisk, Christopher Hertzog, Sankaran N. Nair, Wendy A. Rogers, and Joseph Sharit. 2006. Factors predicting the use of technology: findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). *Psychology and aging* 21, 2 (June 2006), 333–352. <https://doi.org/10.1037/0882-7974.21.2.333> Place: United States.
- [15] Burcu Demiray, Marianne Mischler, and Mike Martin. 2019. Reminiscence in Everyday Conversations: A Naturalistic Observation Study of Older Adults. *The Journals of Gerontology: Series B* 74, 5 (June 2019), 745–755. <https://doi.org/10.1093/geronb/gbx141>
- [16] Yang Deng, Wenxuan Zhang, Yifei Yuan, and Wai Lam. 2023. Knowledge-enhanced Mixed-initiative Dialogue System for Emotional Support Conversations. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Anna Rogers, Jordan Boyd-Graber, and Naoko Okazaki (Eds.). Association for Computational Linguistics, Toronto, Canada, 4079–4095. <https://doi.org/10.18653/v1/2023.acl-long.225>
- [17] Smit Desai and Michael Twidale. 2023. Metaphors in Voice User Interfaces: A Slippery Fish. *ACM Transactions on Computer-Human Interaction* 30, 6 (Dec. 2023), 1–37. <https://doi.org/10.1145/3609326>
- [18] Edwinn Gamborino, Alberto Herrera Ruiz, Jing-Fen Wang, Tsung-Yuan Tseng, Su-Ling Yeh, and Li-Chen Fu. 2021. Towards Effective Robot-Assisted Photo Reminiscence: Personalizing Interactions Through Visual Understanding and Inferring. In *Cross-Cultural Design. Applications in Cultural Heritage, Tourism, Autonomous Vehicles, and Intelligent Agents*, Pei-Luen Patrick Rau (Ed.). Springer International Publishing, Cham, 335–349.
- [19] D Greenberg, M Eacott, D Brechin, and D Rubin. 2005. Visual memory loss and autobiographical amnesia: a case study. *Neuropsychologia* 43, 10 (2005), 1493–1502. <https://doi.org/10.1016/j.neuropsychologia.2004.12.009>
- [20] Kenta Hama, Atsushi Otsuka, and Ryo Ishii. 2024. Emotion Recognition in Conversation with Multi-step Prompting Using Large Language Model. In *Social Computing and Social Media*, Adela Coman and Simona Vasilache (Eds.). Springer Nature Switzerland, Cham, 338–346.
- [21] Jun Li Jeung and Janet Yi-Ching Huang. 2024. Unlocking Memories with AI: Exploring the Role of AI-Generated Cues in Personal Reminiscing. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–6. <https://doi.org/10.1145/3613905.3650979>
- [22] Yucheng Jin, Wanling Cai, Li Chen, Yizhe Zhang, Gavin Doherty, and Tonglin Jiang. 2024. Exploring the Design of Generative AI in Supporting Music-based Reminiscence for Older Adults. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–17. <https://doi.org/10.1145/3613904.3642800>

- [23] Bumsu Kang, Seungwoo Kang, and Inseok Hwang. 2021. MomentMeld: AI-augmented Mobile Photographic Memento towards Mutually Stimulatory Inter-generational Interaction. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–16. <https://doi.org/10.1145/3411764.3445688>
- [24] Masatomo Kobayashi, Akihiro Kosugi, Hironobu Takagi, Miyuki Nemoto, Kiyotaka Nemoto, Tetsuaki Arai, and Yasunori Yamada. 2019. Effects of Age-Related Cognitive Decline on Elderly User Interactions with Voice-Based Dialogue Systems. In *Human-Computer Interaction – INTERACT 2019*, David Lamas, Fernanda Loizides, Lennart Nacke, Helen Petrie, Marco Winckler, and Panayiotis Zaphiris (Eds.). Springer International Publishing, Cham, 53–74.
- [25] Takeshi Kojima, Shixiang (Shane) Gu, Machel Reid, Yutaka Matsuo, and Yusuke Iwasawa. 2022. Large Language Models are Zero-Shot Reasoners. In *Advances in Neural Information Processing Systems*, S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh (Eds.), Vol. 35. Curran Associates, Inc., 22199–22213. [https://proceedings.neurips.cc/paper\\_files/paper/2022/file/8bb0d291acd4acf06ef12099c16f326-Paper-Conference.pdf](https://proceedings.neurips.cc/paper_files/paper/2022/file/8bb0d291acd4acf06ef12099c16f326-Paper-Conference.pdf)
- [26] Emily Kuang, Minghao Li, Mingming Fan, and Kristen Shinohara. 2024. Enhancing UX Evaluation Through Collaboration with Conversational AI Assistants: Effects of Proactive Dialogue and Timing. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–16. <https://doi.org/10.1145/3613904.3642168>
- [27] Wenqiang Lei, Yao Zhang, Feifan Song, Hongru Liang, Jiaxin Mao, Jiancheng Lv, Zhenglu Yang, and Tat-Seng Chua. 2022. Interacting with Non-Cooperative User: A New Paradigm for Proactive Dialogue Policy. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval* (Madrid, Spain) (SIGIR '22). Association for Computing Machinery, New York, NY, USA, 212–222. <https://doi.org/10.1145/3477495.3532001>
- [28] Zisu Li, Li Feng, Chen Liang, Yuru Huang, and Mingming Fan. 2023. Exploring the Opportunities of AR for Enriching Storytelling with Family Photos between Grandparents and Grandchildren. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 3 (Sept. 2023), 1–26. <https://doi.org/10.1145/3610903>
- [29] Amna Liaqat, Benett Axtell, and Cosmin Munteanu. 2022. "With a hint she will remember": Collaborative Storytelling and Culture Sharing between Immigrant Grandparents and Grandchildren Via Magic Thing Designs. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 268 (nov 2022), 37 pages. <https://doi.org/10.1145/3555158>
- [30] Siân E. Lindley. 2012. Before I Forget: From Personal Memory to Family History. *Human-Computer Interaction* 27, 1–2 (2012), 13–36. <https://doi.org/10.1080/07370024.2012.656065>
- [31] Dan P. McAdams, Jeffrey Reynolds, Martha Lewis, Allison H. Patten, and Phillip J. Bowman. 2001. When Bad Things Turn Good and Good Things Turn Bad: Sequences of Redemption and Contamination in Life Narrative and their Relation to Psychosocial Adaptation in Midlife Adults and in Students. *Personality and Social Psychology Bulletin* 27, 4 (April 2001), 474–485. <https://doi.org/10.1177/0146167201274008>
- [32] Anne Marie Piper, Nadir Weibel, and James Hollan. 2013. Audio-enhanced paper photos: encouraging social interaction at age 105. In *Proceedings of the 2013 conference on Computer supported cooperative work*. ACM, San Antonio Texas USA, 215–224. <https://doi.org/10.1145/2441776.2441802>
- [33] Yunfan Shao, Linyang Li, Junqi Dai, and Xipeng Qiu. 2023. Character-LLM: A Trainable Agent for Role-Playing. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, Houda Bouamor, Juan Pino, and Kalika Bali (Eds.). Association for Computational Linguistics, Singapore, 13153–13187. <https://aclanthology.org/2023.emnlp-main.814>
- [34] Gabriel Skantze. 2021. Turn-taking in Conversational Systems and Human-Robot Interaction: A Review. *Computer Speech & Language* 67 (2021), 101178. <https://doi.org/10.1016/j.csl.2020.101178>
- [35] Wilson Tam, Sum Nok Poon, Rathika Mahendran, Ee Heok Kua, and Xi Vivien Wu. 2021. The effectiveness of reminiscence-based intervention on improving psychological well-being in cognitively intact older adults: A systematic review and meta-analysis. *International Journal of Nursing Studies* 114 (Feb. 2021), 103847. <https://doi.org/10.1016/j.ijnurstu.2020.103847>
- [36] Karen P. Tang, Jason I. Hong, Ian E. Smith, Annie Ha, and Lalatendu Satpathy. 2007. Memory karaoke: using a location-aware mobile reminiscence tool to support aging in place. In *Proceedings of the 9th international conference on Human computer interaction with mobile devices and services*. ACM, Singapore, 305–312. <https://doi.org/10.1145/1377999.1378023>
- [37] Elizabeth Thiry, Siân Lindley, Richard Banks, and Tim Regan. 2013. Authoring personal histories: exploring the timeline as a framework for meaning making. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Paris France, 1619–1628. <https://doi.org/10.1145/2470654.2466215>
- [38] Elizabeth Thiry and Mary Beth Rosson. 2012. Unearthing the family gems: design requirements for a digital reminiscing system for older adults. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems*. ACM, Austin Texas USA, 1715–1720. <https://doi.org/10.1145/2212776.2223698>
- [39] Seiki Tokunaga, Kazuhiro Tamura, and Mioko Otake-Matsuura. 2021. A Dialogue-Based System with Photo and Storytelling for Older Adults: Toward Daily Cognitive Training. *Frontiers in Robotics and AI* 8 (June 2021), 644964. <https://doi.org/10.3389/frobt.2021.644964>
- [40] Xin Wang, Juan Li, Tianyi Liang, Wordh Ul Hasan, Kimia Tuz Zaman, Yang Du, Bo Xie, and Cui Tao. 2024. Promoting Personalized Reminiscence Among Cognitively Intact Older Adults Through an AI-Driven Interactive Multimodal Photo Album: Development and Usability Study. *JMIR aging* 7 (Jan. 2024), e49415. <https://doi.org/10.2196/49415>
- [41] J. D. Webster. 1993. Construction and Validation of the Reminiscence Functions Scale. *Journal of Gerontology* 48, 5 (Sept. 1993), P256–P262. <https://doi.org/10.1093/geronj/48.5.P256>
- [42] Xiaoying Wei, Yizheng Gu, Emily Kuang, Xian Wang, Beiyian Cao, Xiaofu Jin, and Mingming Fan. 2023. Bridging the Generational Gap: Exploring How Virtual Reality Supports Remote Communication Between Grandparents and Grandchildren. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) (CHI '23). Association for Computing Machinery, New York, NY, USA, Article 444, 15 pages. <https://doi.org/10.1145/3544548.3581405>
- [43] Gerben J. Westerhof and Ernst T. Bohlmeijer. 2014. Celebrating fifty years of research and applications in reminiscence and life review: State of the art and new directions. *Journal of Aging Studies* 29 (April 2014), 107–114. <https://doi.org/10.1016/j.jaging.2014.02.003>
- [44] Tim Wildschut, Constantine Sedikides, Jamie Arndt, and Clay Routledge. 2006. Nostalgia: Content, triggers, functions. *Journal of Personality and Social Psychology* 91, 5 (Nov. 2006), 975–993. <https://doi.org/10.1037/0022-3514.91.5.975>
- [45] Shuai Yang, Yuying Ge, Yang Li, Yukang Chen, Yixiao Ge, Ying Shan, and Yingcong Chen. 2024. SEED-Story: Multimodal Long Story Generation with Large Language Model. arXiv:2407.08683 [cs.CV]. <https://arxiv.org/abs/2407.08683>
- [46] Shukang Yin, Chaoyou Fu, Sirui Zhao, Ke Li, Xing Sun, Tong Xu, and Enhong Chen. 2024. A Survey on Multimodal Large Language Models. arXiv:2306.13549 [cs.CV]. <https://arxiv.org/abs/2306.13549>
- [47] Chuanqi Zang, Jiji Tang, Rongsheng Zhang, Zeng Zhao, Tangjie Lv, Mingtao Pei, and Wei Liang. 2024. Let Storytelling Tell Vivid Stories: An Expressive and Fluent Multimodal Storyteller. arXiv:2403.07301 [cs.CV]. <https://arxiv.org/abs/2403.07301>
- [48] Yu Zhang, Jingwei Sun, Li Feng, Cen Yao, Mingming Fan, Liuxin Zhang, Qianying Wang, Xin Geng, and Yong Rui. 2024. See Widely, Think Wisely: Toward Designing a Generative Multi-agent System to Burst Filter Bubbles. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '24). Association for Computing Machinery, New York, NY, USA, Article 484, 24 pages. <https://doi.org/10.1145/3613904.3642545>
- [49] Zhongyu Zhang, Lina Xu, Xingkai Wang, Xu Zhang, and Mingming Fan. 2025. Understanding and Co-designing Photo-based Reminiscence with Older Adults. *Proc. ACM Hum.-Comput. Interact.* 9, 2, Article CSCW196 (April 2025), 30 pages. <https://doi.org/10.1145/3711094>
- [50] Qingxiao Zheng, Yiliu Tang, Yiren Liu, Weizi Liu, and Yun Huang. 2022. UX Research on Conversational Human-AI Interaction: A Literature Review of the ACM Digital Library. In *CHI Conference on Human Factors in Computing Systems*. ACM, New Orleans LA USA, 1–24. <https://doi.org/10.1145/3491102.3501855>

## A Prompts for Artifact Descriptions and Examples

### A.1 Prompt for Generating Personal Photo Descriptions and an Example

Prompt for generating description for a personal photo:

Describe this personal photo, including the people, objects, atmosphere, and expressions. Also, provide potential keywords that might appear in a conversation reminiscing about the past based on this photo. Do not use line breaks; instead, use \n to indicate a new line.

Example of description of a personal photo:

This black and white photo shows a middle-aged couple sitting in a living room.\nIn the center of the photo is an old-style television set with a vase on top.\nDecorative paintings and a calendar are hanging on the wall.\n\nThe room also contains potted plants and other furniture.\n\nThe couple is wearing neat white

shirts and has calm expressions.\nThe overall atmosphere is warm and comfortable, reflecting the family life of that time.\nPotential keywords:\nOld photo, 1960s-70s, husband and wife, living room, old-style TV, family life, nostalgia, memories, younger days, new household appliances, interior design style, changes in lifestyle.

## A.2 Prompt for Generating Generic Nostalgic Object Descriptions and an Example

Prompt for generating description for a generic nostalgic object:

Describe this 3D model of generic nostalgic object, including its appearance, elements, historical background, function, and everyday uses. Also, provide potential keywords that might appear in a conversation reminiscing about the past based on this object. Do not use line breaks; instead, use \n to indicate a new line.

Example of description of a generic nostalgic object (TV as an example):

This 3D model is an old-style television set.\nAppearance: Square wooden cabinet, large screen on the front, two knobs and a speaker grille on the right side.\nElements: Cathode ray tube screen, tuning knob, volume knob, wooden casing, metal decorations.\nHistorical background: This type of TV was common in the 1960s-70s and was the center of home entertainment.\nFunction: Receives television signals and displays images and sound.\nEveryday uses: Watching news, TV shows, movies, etc.; families gathering around the TV to spend time together.\nPotential keywords: Childhood memories, family gatherings, black and white TV, adjusting antennas, static screen, limited broadcasting hours, weekly must-watch programs, TV repairman.

## B Prompts for Artifact Switching Decision

Prompts for *Artifact Switching Agent*'s decision making:

I (an agent) am reminiscing about the past with an elderly person (human) based on old objects (3D models) and their old photos. You, as an invisible assistant in our conversation, should assist our dialogue according to the following rules:

1. If the human's latest dialogue mentions an image more strongly related than the current one, tell me to change the topic and select a matching image ID from the image list.
2. If the elderly person cannot engage in the conversation with relevant personal experiences, for example, if they reply very briefly, express little to no experience with the items in the image, mainly talk about others' experiences, or show disinterest, tell me to change the topic and select an image ID from the list that the human might be interested in.
3. Check how many rounds of conversation have occurred around the current 3D model or photo. If it

exceeds 2 rounds for a 3D model or 5 rounds for a photo, tell me to change the topic and select an image ID from the list.

4. When selecting an image from the list, if the current image is a 3D model, prioritize selecting a personal photo; if the current image is a personal photo, prioritize selecting a 3D model.
5. Output in JSON format, without any explanation.

—

Conversation:

`<conversation_info>`

—

Image list, each line in the format 'image id , image name , brief description of image':

`<image_info>`

—

JSON output format:

`"change_topic": true, # or false`

`"image_id": 2, # the id of the image, leave -1 if 'change_topic' is false`

Your response:

What the *Artifact Switching Agent* (assistant) say in conversation:

"I have found this relevant image: `<image_info><model_change>`.

If this is a description of a photo, please be aware that this is a personal photo of the human. Guide the human to reminisce about the scene in the photo, including time, place, people, background, etc., and encourage storytelling. If this is a description of a 3D model of an old object, the human will be viewing this model in real-time. Please naturally encourage the elderly person to observe the model's details, combining its appearance and elements (for example: 'Do you see [appearance or element]...'). If the human has already discussed some content in previous conversations, please inquire about aspects of the image that haven't been mentioned yet, avoiding repetition. After briefly summarizing the previous topic, you now need to smoothly transition the conversation to this new image."

## C Prompts for Responding Agents

Agent role setting for the older agent (for prompt constant `<role_prompt>`):

You are agent\_1. You play the role of a 65-75 year old Chinese elderly person, similar in age to the human. You are gentle and patient, helping the human reminisce about good times in the past, and encouraging the human to share their own stories and feelings.

Agent role setting for the younger agent (for prompt constant `<role_prompt>`):

You are agent\_2. You play the role of a 12-year-old Chinese child. You are full of energy and use simple, direct language to guide the human in sharing their experiences and feelings. Your language should make the human feel relaxed and happy.

The prompt for generating responses are:

You are reminiscing about the past with an elderly person (human), and this is your character setting: <role\_prompt>. The human's gender is <user\_gender>. Please speak in Chinese. Your response should be a short paragraph, not too long, without line breaks. Avoid using questions, and instead use your own personal experiences as guidance. If you must use a question, limit it to only one. In the conversation, besides you, there is another agent and a conversation assistant (assistant). The assistant will provide you with image descriptions and topic guidance. Please do not engage in any form of dialogue with the other agent or the assistant. Please continue the conversation with the human about reminiscing the past. If the human asks you a question, please answer their question first.

## D Post-study Survey

The post-study survey instrument is listed in Figure 15.

## E Semi-structured Interview Outline

The Semi-structured interview outline is listed as follows:

### (a) Conversational Agents:

- Ratings in the Post-Study Survey
  - What were the reasons behind the ratings you provided in the post-study survey?
- Perceptions and Impact of Agents on Conversational Behavior
  - Who is your favorite conversational character and why?
  - What are your feelings towards different agents? Do they influence your conversational behavior?
- Preferences for Character Design
  - How would your experience differ if these virtual characters were your friends or children?
  - Are there other characters you would like to interact with?
  - What are your expectations and suggestions for character design?

### (b) Nostalgic Artifacts:

- Ratings in the Post-Study Survey
  - What were the reasons behind the ratings you provided in the post-study survey?
- Memorable 3D Nostalgic Objects or Personal Photos
  - Were there any particularly memorable 3D nostalgic objects or personal photos, and what dialogues were associated with them? What made them memorable?
- Differences in Perceptions Between 3D Nostalgic Objects and Personal Photos
  - How do your feelings about 3D nostalgic objects differ from personal photos? Did they affect your conversational behavior?

### (c) Experience During System Use:

- Ratings in the Post-Study Survey
  - What were the reasons behind the ratings you provided in the post-study survey?
- Overall Impression of the Experience
  - What is your overall impression of the experience?

- Memorable Aspects of Using the System
  - What were the most memorable aspects of using the system? What are its greatest strengths and weaknesses?
- Suggestions for System Design
  - Do you have any suggestions for the system design, including dialogue, characters, materials, and other aspects?

### (d) Daily Reminiscence Behavior:

- Habits of Organizing and Reviewing Old Photos
  - Do you have a habit of browsing or organizing old photos?
- Comparison with Daily Reminiscence Behavior
  - How does your daily reminiscence behavior compare with your experience using the system?

Please rate how much you agree or disagree with each statement by choosing one of the options below.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	
1. Overall, I enjoyed talking to the four agents.	<input type="radio"/>					
2.	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	
Overall, I like talking to <i>Zhang Jianguo</i> (the older male agent).	<input type="radio"/>					
Conversations with <i>Zhang Jianguo</i> were natural.	<input type="radio"/>					
Conversations with <i>Zhang Jianguo</i> were smooth.	<input type="radio"/>					
Conversations with <i>Zhang Jianguo</i> were efficient.	<input type="radio"/>					
I believe <i>Zhang Jianguo</i> is trustworthy.	<input type="radio"/>					
3.	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	
Overall, I like talking to <i>Liu Lifen</i> (the older female agent).	<input type="radio"/>					
Conversations with <i>Liu Lifen</i> were natural.	<input type="radio"/>					
Conversations with <i>Liu Lifen</i> were smooth.	<input type="radio"/>					
Conversations with <i>Liu Lifen</i> were efficient.	<input type="radio"/>					
I believe <i>Liu Lifen</i> is trustworthy.	<input type="radio"/>					
4.	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	
Overall, I like talking to <i>Wang Xiaoming</i> (the younger male agent).	<input type="radio"/>					
Conversations with <i>Wang Xiaoming</i> were natural.	<input type="radio"/>					
Conversations with <i>Wang Xiaoming</i> were smooth.	<input type="radio"/>					
Conversations with <i>Wang Xiaoming</i> were efficient.	<input type="radio"/>					
I believe <i>Wang Xiaoming</i> is trustworthy.	<input type="radio"/>					
5.	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	
Overall, I like talking to <i>Chen Xiaohong</i> (the younger female agent).	<input type="radio"/>					
Conversations with <i>Chen Xiaohong</i> were natural.	<input type="radio"/>					
Conversations with <i>Chen Xiaohong</i> were smooth.	<input type="radio"/>					
Conversations with <i>Chen Xiaohong</i> were efficient.	<input type="radio"/>					
I believe <i>Chen Xiaohong</i> is trustworthy.	<input type="radio"/>					
6.	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	
The 3D nostalgic objects appeared in the conversations were appropriate.	<input type="radio"/>					
I enjoyed engaging in conversations about the 3D nostalgic objects.	<input type="radio"/>					
The 3D nostalgic objects stimulated my recall of the past.	<input type="radio"/>					
The 3D nostalgic objects increased my willingness to use this system.	<input type="radio"/>					
7.	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	
My personal photos appeared in the conversations were appropriate.	<input type="radio"/>					
I enjoyed engaging in conversations about my personal photos.	<input type="radio"/>					
My personal photos stimulated my recall of the past.	<input type="radio"/>					
My personal photos increased my willingness to use this system.	<input type="radio"/>					
8. I would like to use this system in my daily life.	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	
9. I feel that my mood has improved after using the system.	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	
10. I didn't feel sick when using the system.	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	

Please select the option that matches your situation.

11. How much time do you spend using these electronic devices each day?  <1 hour  1-3 hours  3-5 hours  >5 hours

12. Have you used intelligent voice assistants? (e.g., XiaoAi Speaker, Siri, XiaoDu Speaker, etc.)  Yes  No

13. How often do you organize or browse old photos?  Never  Rarely  Sometimes  Often  Always

Figure 15: Post-study survey.