



AI as a Bridge Across Ages: Exploring The Opportunities of Artificial Intelligence in Supporting Inter-Generational Communication in Virtual Reality

QIUXIN DU*, Beijing Institute of Technology, China

XIAOYING WEI*, The Hong Kong University of Science and Technology, China

JIawei LI, The Hong Kong University of Science and Technology (Guangzhou), China

EMILY KUANG, Rochester Institute of Technology, USA

JIE HAO, Beijing Institute of Technology, China

DONGDONG WENG, Beijing Institute of Technology, China

MINGMING FAN[†], The Hong Kong University of Science and Technology (Guangzhou), China and The Hong Kong University of Science and Technology, China

Inter-generational communication plays a vital role in bridging generational gaps and fostering mutual understanding. However, it remains challenging due to differences in cultural norms, communication styles, and geographical separation. While prior studies have shown Virtual Reality (VR) as a medium that fosters a relaxed atmosphere and companionship, its capacity to address the nuanced dynamics of inter-generational dialogue, such as divergent values and relational intricacies, remains limited. To address this gap, we explored the opportunities of Artificial Intelligence (AI) to support inter-generational communication in VR. We developed three technology probes (e.g., Content Generator, Communication Facilitator, and Info Assistant) in VR and employed them in a probe-based participatory design study with twelve inter-generational pairs. Our results show that AI-powered VR facilitates inter-generational communication by enhancing mutual understanding, fostering conversation fluency, and promoting active participation. We also introduce several challenges when using AI-powered VR in supporting inter-generational communication and derive design implications for future VR platforms, aiming to improve inter-generational communication.

CCS Concepts: • **Human-centered computing** → **Virtual reality; User studies; Empirical studies in HCI.**

Additional Key Words and Phrases: Virtual Reality; Family; Social Interaction; Participatory Design

ACM Reference Format:

Qiuxin Du, Xiaoying Wei, Jiawei Li, Emily Kuang, Jie Hao, Dongdong Weng, and Mingming Fan. 2025. AI as a Bridge Across Ages: Exploring The Opportunities of Artificial Intelligence in Supporting Inter-Generational

*Both authors contributed equally to this research.

[†]Corresponding Author

Authors' Contact Information: [Qiuxin Du](#), Beijing Institute of Technology, Beijing, China, qiuxindu@bit.edu.cn; [Xiaoying Wei](#), The Hong Kong University of Science and Technology, Hong Kong SAR, China, xweias@connect.ust.hk; [Jiawei Li](#), The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China, crane787213823@outlook.com; [Emily Kuang](#), Rochester Institute of Technology, Rochester, USA, ek8093@rit.edu; [Jie Hao](#), Beijing Institute of Technology, Beijing, China, haoj@bit.edu.cn; [Dongdong Weng](#), Beijing Institute of Technology, Beijing, China, crgj@bit.edu.cn; [Mingming Fan](#), The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China and The Hong Kong University of Science and Technology, Hong Kong SAR, China, mingmingfan@ust.hk.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2025 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM 2573-0142/2025/4-ARTCSW026

<https://doi.org/10.1145/3710924>

Communication in Virtual Reality. *Proc. ACM Hum.-Comput. Interact.* 9, 2, Article CSCW026 (April 2025), 29 pages. <https://doi.org/10.1145/3710924>

1 Introduction

Inter-generational communication refers to the exchange of experiences and perspectives across generations, fostering a deeper mutual understanding and bridging the gap between different generations within a family [45, 96, 105]. Inter-generational communication plays a crucial role in nurturing emotional growth, enhancing self-esteem, and cementing familial bonds for younger members [11, 98, 106], while simultaneously providing older generations with essential social engagement and a cherished role in the family structure [17, 18, 59, 77, 89]. However, modern society's rapid technological and cultural shifts pose considerable challenges to maintaining effective inter-generational communication. The disparity in cultural backgrounds, communication styles, and subtle emotional expressions, coupled with geographical distances, often hamper the flow and depth of these interactions [93, 100].

Recent studies have seen the emergence of virtual reality (VR) as a promising medium to support inter-generational communication. This platform offers unique affordances, such as customizable avatars, immersive virtual environments, and a sense of co-presence, which have shown promise in fostering more relaxed, equal, and emotionally fulfilling inter-generational interactions [5, 36, 54, 102]. Studies also showed that communication via VR could improve the overall quality of life for older people while decreasing the caregiver burden of their younger generations [2, 36]. While previous research has explored various aspects of VR, they often focused on predefined features and specific usage contexts [2, 36, 56, 102], not fully addressing the complexities of inter-generational dialogue, such as differing values, generational hierarchy, evolving emotional states, or relational dynamics (e.g., conflict management) [21, 100].

Previous research has shown artificial intelligence's (AI) potential to bridge these gaps by understanding and responding to users' dialogue content and emotions in real-time [8, 112]. This includes technologies like large language models (LLMs), which can comprehend and generate human-like text, and AI-generated content (AIGC), which involves creating various types of media using AI. Researchers leveraged these technologies to support inter-generational communication in non-immersive environments [39, 43, 44, 111], revealing AI's ability to manage complex communication dynamics and promote reminiscence [43, 44, 111, 112]. However, there is limited understanding of how AI can be used in VR to support inter-generational communication. VR offers a flexible platform for AI to present diverse outputs (e.g., text, audio, images, 3D models) through interactive avatars and immersive environments, while allowing users to physically interact with AI-generated objects [13, 31, 33, 66, 88], potentially transforming inter-generational interactions. To understand how inter-generational pairs would use AI in VR to facilitate communication, we propose our first research question, **RQ1: How could AI-powered VR be leveraged to support inter-generational communication?**

Furthermore, previous works have shown several challenges in using AI for interpersonal communication. For instance, AI interventions could disrupt the smooth flow of group discussions [24, 55, 64]. Unlike formal communication, inter-generational communication within families tends to be more casual and focused on personal topics, often reflecting familial backgrounds [34, 58]. This informal and personal communication style could present new challenges for AI in understanding and appropriately intervening in inter-generational exchanges. This leads us to explore our second research question: **RQ2: What are the challenges of using AI to support VR-based inter-generational communication?**

Building on previous research that demonstrated the effectiveness of technology probe-based participatory design in eliciting user feedback and identifying opportunities in emerging technologies [42, 56, 81], we adopted a similar methodology. Using three types of AI probes (e.g., Content Generator, Communication Facilitator, and Info Assistant), we conducted a probe-based participatory design study with twelve inter-generational pairs to uncover the effectiveness and user expectations of AI-powered VR in supporting communication. Participants first tried out these AI probes to inspire them to think about the possible ways in which AI could enhance their communication in VR. They then participated in a co-design workshop to brainstorm and design desired AI features to assist their communication.

Our findings illustrate the multifaceted advantages of AI-powered VR in supporting inter-generational communication. Drawing from the outcomes of our observation and participatory design workshop, participants leveraged AI-powered VR to 1) foster mutual understanding by transforming abstract conversation content into concrete forms for better idea exchange, along with offering various ways to interact with the generated content to deepen communication; 2) enhance conversation fluency by personalizing AI representations for better acceptance, and providing interventions to prevent awkward silences, maintain balanced dialogue, and manage potential frictions; 3) promote their active participation by modifying users' avatars or virtual environments based on their emotions and contextual contents. Our results also reveal the challenges of using AI-powered VR to support inter-generational communication. Based on the results, we proposed design implications for future designers to better leverage AI to support VR-based communication.

In sum, we make the following contributions:

- Our study advances the understanding of AI-powered VR in promoting inter-generational communication. We demonstrate AI's effectiveness in enhancing mutual understanding, improving conversational fluency, and encouraging active participation, highlighting its role in bridging generational gaps in VR interactions.
- Drawing from participatory design outcomes, we outline participants' insights into customizing AI features to address specific communication needs in inter-generational interactions, as well as provide design implications for future researchers and developers to create more practical and meaningful AI-based VR platforms.

2 Related Work

2.1 Benefits and challenges of inter-generational communication within a family

Inter-generational communication involves exchanging information, beliefs, and feelings between individuals from different generations. Inter-generational communication within families (i.e., connection across generations like grandparents, parents, and children) is particularly beneficial [45, 96, 105]. On the one hand, it supports the growth and emotional development of younger generations, enhancing their self-esteem and deepening familial bonds [11, 98, 106]. On the other hand, it provide older generations with valuable social engagement and reducing feelings of isolation [17, 59, 77, 89]. This helps maintain an active and cherished role within the family dynamic.

However, maintaining successful inter-generational communication is challenging in today's society, which is characterized by rapid technological and cultural shifts. First, while people have a strong need and desire to stay connected with their families to share recent life events, two generations find it *difficult to achieve mutual understanding* due to disparate cultural backgrounds, including distinct values, beliefs, and communication styles [73, 97]. This often leads to misinterpretations and conflicts in their exchanges [75, 84]. Second, the generational hierarchy often results in the younger generation demonstrating *over-accommodations* towards the elderly [46], such as listening passively without interrupting. Although these actions are well-intentioned and display

respect for the older generation, they frequently lead to miscommunication, a lack of genuine connection, and frustration among young people [32, 104]. Third, *the subtlety of emotional expression* between two generations hinders successful inter-generational communication, especially in countries with Eastern cultures [48, 94].

Additionally, geographical distance between family members often impedes intimate inter-generational communication [26, 34]. While audio-visual tools like audio or video calls enable remote connections, they do not always provide enough emotional support to give a feeling of closeness and intimacy. Therefore, further study on how to facilitate more active and effective inter-generational communication is warranted by HCI researchers.

2.2 VR for supporting inter-generational communication

Previous works in the HCI field have leveraged VR to support inter-generational communication [2, 36, 56, 102]. These studies demonstrated three affordances of VR that can support communication between two generations. First, VR allows users to represent themselves via customizable avatars. Such representation affects their perceptions of each other and blurs the generation gaps, thereby fostering a more casual and equal communication experience than in reality [102]. Additionally, the study indicated that users could engage in various activities in a virtual environment, which stimulates more topics for them to keep their conversations flowing [2, 36, 101, 102]. Finally, VR's immersive environment offers the sense of co-presence for two generations to achieve a degree of companionship [1, 54], which is important for maintaining close emotional connections. Researchers also showed that using VR as a long-term communication tool improved the overall quality of life for older people while decreasing the caregiver burden of younger generations [2].

While previous studies have demonstrated the potential of VR in facilitating inter-generational communication, their emphasis has predominantly been on exploring predefined features to meet user needs within specific communicative contexts. For instance, Wang et al. [101] and Hoeg et al. [36] explored the design of collaborative activities (e.g., tandem biking, co-learning) in VR for two generations. Wei et al. investigated the design of avatar appearances in VR to facilitate their communication [102]. However, these predefined designs fall short of addressing the myriad complexities inherent in inter-generational dialogue, such as navigating differing knowledge, over-accommodation, and the evolving nature of users' conversational content and emotional states [32, 73, 97, 104]. Therefore, this work explores how to design an interactive VR environment that effectively addresses the challenges present during inter-generational communication in VR.

2.3 Integrating AI into inter-generational communication within VR environments

Previous research has demonstrated the potential of AI to moderate inter-generational communication by understanding dialogue's content and emotions in real-time, and intervening accordingly [8, 112]. Studies have leveraged AI to support inter-generational communication in non-immersive environments, including providing context-related photos to arouse reminiscence and foster conversation [43], generating story-related questions to enhance learning outcomes [25, 111], and answering users' questions to improve the collaboration efficacy [53].

While AI shows potential in enhancing inter-generational communication, its integration into immersive VR environments remains under-explored. The integration of AI and VR has the potential to facilitate inter-generational communication for two main reasons. First, AI could enable VR to dynamically adjust its environment and settings by monitoring users' dialogue in real-time, potentially addressing the limitations associated with predefined designs in previous VR research [24, 55, 61, 112]. Second, unlike non-immersive tools that offer limited interaction (e.g., text-based or speech-based agents [112]), immersive VR environments present a unique opportunity

to leverage various advanced AI outputs to enhance communication, including interactive and responsive virtual avatars, dynamic 3D visualizations, and adaptable virtual spaces [13, 31, 33, 66, 88]. This combination could potentially address the inter-generational communication challenges presented in Section 2.1, such as providing various AI-generated content (e.g., text, images, objects) to exchange ideas for achieving mutual understanding.

Despite these potential benefits, little is known about how inter-generational pairs perceive and leverage AI in VR to support communication, as well as the challenges and considerations in designing AI features for better supporting inter-generational communication. Understanding these aspects could help future researchers and developers create more practical and meaningful AI-based VR platforms. To address these gaps, we employed a probe-based participatory design approach to explore the effectiveness and user perceptions of AI in enhancing inter-generational communication in VR.

2.4 Technology probe methodology

A technology probe study is an innovative research methodology primarily used in the fields of HCI and CSCW [37, 38, 56]. This method typically offers fundamental functionalities of specific emerging technology, enabling users to explore and understand their interactions with these innovations [38]. The primary purpose of a technology probe study is to inspire and provoke responses from users, shedding light on user needs, desires, and technological possibilities, rather than to test a specific hypothesis or to refine a product for real-life use. This is particularly valuable in the early stages of technology development, where understanding the user experience is crucial [56, 81].

Prior work has shown the feasibility of the technology probe study in identifying opportunities of emerging technology in inter-generational communication, such as AR [56] and VR [102]. By observing how participants use and adapt to the technology, these studies identify unforeseen opportunities and challenges, leading to more user-centered design decisions. Therefore, in this work, we adopted the technology probe study method to explore AI's possibilities for facilitating inter-generational communication in immersive environments and answering our RQs.

3 Method

To answer our RQ, we designed and implemented AI probes, and then conducted a probe-based participatory design workshop to elicit participants' perceptions and envisions. For simplicity, in the rest of this paper, the older and younger generation are abbreviated as OG and YG respectively (e.g., P1-YG means the younger generation in pair 1).

3.1 Probe Design and Implementation

We developed three types of AI probes, including **Content Generation** (i.e., image generation, 3D object generation, and scene generation), **Communication Facilitator** (i.e., chat facilitator and emotional visualization), and **Info Assistant**, as shown in Figure 1. The probes include three types of AI outputs: text/voice, images, and 3D objects/environments, and can be activated in two ways: passively and proactively. We provide these probes to offer participants fundamental and comprehensive AI functionalities to experience in VR [38], which allows us to understand how inter-generational pairs interact with AI while communicating in VR, what challenges they may encounter, and how they envision future AI features to be designed in VR to better support their communication. In the following section, we introduce the design considerations and implementation details of each AI probe.

3.1.1 Content Generation. As discussed in Section 2.3, VR's capability to showcase various AI outputs enables users to visualize their conversation content and interact with it, potentially

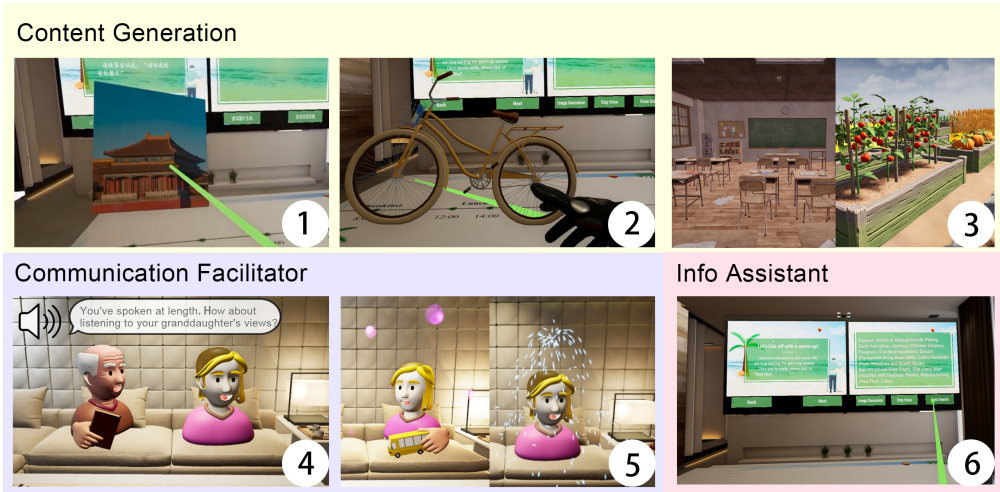


Fig. 1. Illustrations of our AI probes. (1) Image generation: it generates images based on verbal descriptions provided by users. (2) 3D Object generation: it creates virtual 3D models according to users' verbal descriptions. (3) Scene generation: the virtual environment automatically switches based on the content of the conversation. (4) Chat facilitator: it intervenes verbally to balance conversations when one party dominates, and offers potential conversation topics through voice prompts during awkward silences. (5) Emotion visualization: analyzing the emotional tone of chat content, it displays bubble effects for positive emotions (left) and rain effects for negative emotions (right). (6) Info assistant: users press the button to input voice-based queries and receive spoken and text answers in response to their inquiries.

enhancing mutual understanding and bridging knowledge gaps across generations [4]. To explore how inter-generational pairs leverage Content Generation in VR to facilitate communication, we present three fundamental forms of AI outputs to stimulate users' imagination and insight.

Image generation. Images can effectively convey information such as emotions, concepts, and events, playing a crucial role in facilitating communication in daily life [6, 43]. A previous study demonstrated that sharing and discussing images in a VR environment provided higher feelings of co-presence and increased the quality of interaction compared to screen-based communication (e.g., Skype) [54]. In this work, we propose an Image generation function in VR. This feature is activated when users click a button and provide verbal descriptions of an image. The system then generates the image based on these voice prompts and displays it in front of two users. Users can manipulate these images by grabbing, moving, zooming, or rotating with their hands, as illustrated in the first image of Figure 1. If the generated images do not meet the users' expectations, they can be discarded and dragged into a virtual trash bin. For implementation, we used Azure AI Speech for voice transcription¹ and DALL-E2 for image generation².

3D object generation. 3D objects are particularly useful for explaining specific objects or concepts [23]. Considering the rendering time and computing costs of 3D object generation from text, we opted to provide users with a selection of pre-generated 3D objects relevant to the tasks, such as various vehicles (bus, taxi, etc.) for the first task and assorted stationery items (vintage chair, ruler, pencil case, printer, etc.) for the second, as shown in Figure 5. For implementation, we used Point-E to generate 3D models from text [76]. Additionally, we implemented a keyword detection that automatically activates this feature whenever users mention specific terms from our

¹<https://azure.microsoft.com/en-us/products/ai-services/ai-speech>

²<https://openai.com/dall-e-2>

curated list. Upon detection, these 3D objects materialize in front of the users. Users can interact with these 3D objects by grabbing, moving, zooming, or rotating them.

Scene generation. A content-specific environment creates an engaging and active communication atmosphere for inter-generational users [102]. Similar to 3D object generation, we provided pre-created virtual scenes pertinent to the tasks, such as a living room or a farm for the first task and both a new and an old classroom for the second, as shown in Figure 6 in the Appendix. These scenes are activated through keyword recognition. The scenes were developed using Set-the-Scene [16].

3.1.2 Communication Facilitator. Integrating an AI coordinator to monitor dialogue and offer timely interventions enhances communication efficiency [24, 55, 112]. To explore the possibilities of VR to support AI coordinators in inter-generational contexts, we developed the chat facilitator (voice-based intervention) and emotional visualization (visual-based intervention).

Chat Facilitator. As discussed in Section 2.3, *over-accommodation* and *the lack of conversation topic* often hinder effective inter-generational communication. Therefore, we provided verbal interventions. We detected *imbalances* (i.e., one party speaks continuously for over one minute while the other remains silent) and *silences* (i.e., no one speaks for more than thirty seconds) by recording the length of silence in real-time, and utilized prompt engineering [49] with ERNIE Bot³ to provide appropriate voice suggestions. For example, if one party dominates the conversation, the Chat Facilitator might suggest, “How about listening to the other’s opinions?”

Emotional Visualization. We also provide visual interventions in VR to help inter-generational users in *expressing their emotions more effectively*. We employed keyword monitoring to detect positive (e.g., ‘happy’, ‘excited’) and negative emotions (e.g., ‘sad’, ‘tired’). Subsequently, we incorporated specific visual effects that corresponded with these emotions. For example, the rain effect was chosen to express their negative emotions, as rain is usually associated with sadness. Conversely, bubble effects were used for positive emotions, reflecting their joyful connotations, as shown in the fifth image in Figure 1.

3.1.3 Info Assistant. Prior studies showed that people often queried information with AI to support interpersonal communication, helping them communicate more confidently and accurately [24, 27].

Info Assistant. This feature serves two main purposes: (1) to help users easily query information within the VR environment, supporting their daily conversations and experimental tasks; and (2) to allow users to experience voice interactions with AI in VR, inspiring them to envision how AI voice assistants can be better designed in VR, including interaction methods and AI representation. Users can click a button and verbally query the AI, with responses displayed as text on the screen and simultaneously vocalized using Azure AI Speech, as depicted in the sixth image of Figure 1. Participants can halt the audio playback at any time by activating the “Stop Speech” button. The search engine functionality is powered by ChatGPT-3.5⁴.

3.2 System Design and Implementation

Our goal was to integrate all the AI probes into a user-friendly VR application. We developed the system using Unreal Engine 5.1.1⁵. Our system supports remote VR communication through UDP (User Datagram Protocol) for real-time interaction and synchronization of virtual items and scenes across both users’ interfaces. Additionally, we equipped the system with a camera to document participants’ behaviors and interactions within the VR environment.

³<https://yiyan.baidu.com/>

⁴<https://chat.openai.com/>

⁵<https://www.unrealengine.com/en-US/>

For the VR environment, we selected a living room as the task setting, complete with a sofa, coffee table, and typical household decorations, as illustrated in Figure 6 in the Appendix. This cozy environment was chosen to provide users with a familiar and relaxing atmosphere to facilitate their communication [103]. We set up two virtual screens to display different information. The left screen displays text instructions to 1) help users better learn the AI probe during the training phase and 2) guide them through each task during the study session. The right screen offers an interactive interface for 1) invoking Content Generation and Info Assistance functions, and 2) providing a display area for the results of Info Assistance, making it easily viewable for users.

For the avatars, we provided users with four types of avatars to embody in VR: an elderly male, an elderly female, a young male, and a young female, as shown in Figure 7 in the Appendix. Users were required to select one avatar that best matched their real-life identity to mitigate novelty effects [69, 83]. We did not offer an avatar customization function to mitigate distractions related to avatar appearance and focus on interacting with AI in VR. Additionally, to foster an active and relaxing atmosphere that encourages user interaction, we offered abstract and cartoon-style avatars instead of realistic ones [102].

3.3 Participants

We advertised our recruitment information through snowball sampling, social media, and community centers. Before the study, participants completed an online recruitment survey. Based on their responses, we invited participants with diverse backgrounds (e.g., education level, occupation) to participate in the study. We ensured all participants met the inclusion criteria, which were no major medical illnesses, no history of mental or neurological abnormalities, and no sensory impairments. The university’s Institutional Review Board (IRB) approved this study for research ethics.

We recruited 12 pairs of participants (24 in total) from two different regions in China (four pairs were from the metropolitan center, while eight were from regional locations). The YG participants were aged 15 to 30 ($M = 20.6$, $SD = 5.2$), and the OG participants were aged 55 to 78 ($M = 63.2$, $SD = 6.0$). The OG and YG participants in each pair had lived separately for more than one year and did not currently live together. Each pair had diverse relationships (6 pairs were grandparent-grandchild, 2 pairs were parent-child, and 4 pairs were other forms of kinship, such as uncle-niece). Most participants (OG: $n=9$, YG: $n=12$) had experience using AI-based products, including voice assistants (e.g., Siri or Tmall Genie), and generative AI applications (e.g., ChatGPT or ERNIE Bot). Most participants (OG: $n=11$, YG: $n=9$) had no prior experience with VR. Detailed participant background information is shown in Table 1.

Table 1. Demographic information for all participants

Pair ID	Relationship	Age (OG-YG)	VR experience (OG-YG)	AI experience (OG-YG)
1	GP-GC	61-15	No-No	No-Yes
2	kinship	61-22	No-Yes	Yes-Yes
3	kinship	60-25	Yes-No	Yes-Yes
4	Parent-Child	55-20	No-No	Yes-Yes
5	Parent-Child	60-30	No-No	Yes-Yes
6	GP-GC	78-26	No-No	No-Yes
7	kinship	60-25	No-No	Yes-Yes
8	GP-GC	63-15	No-No	No-Yes
9	GP-GC	68-16	No-Yes	Yes-Yes
10	kinship	61-22	No-No	Yes-Yes
11	GP-GC	62-16	No-Yes	Yes-Yes
12	GP-GC	70-15	No-No	Yes-Yes

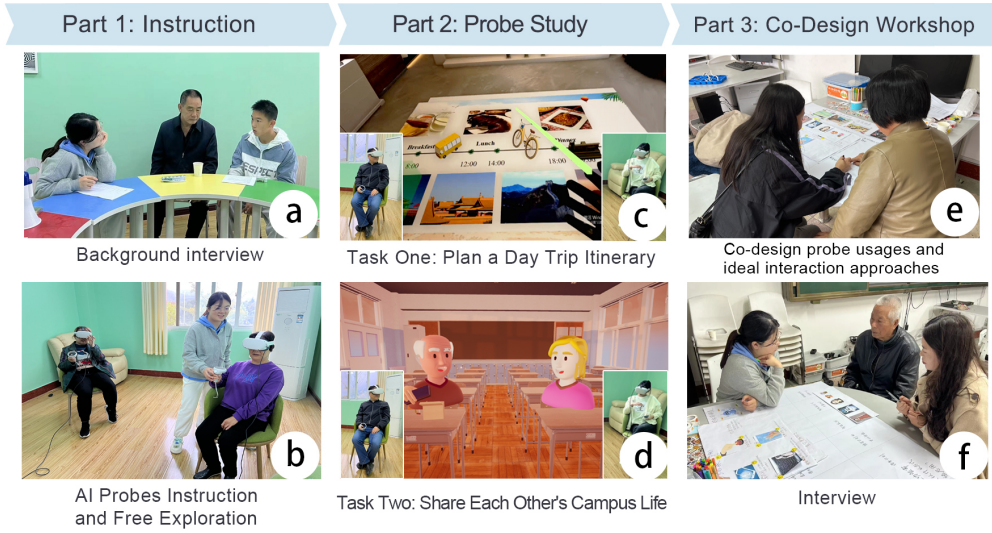


Fig. 2. The user study process is structured into distinct parts, each represented by a column. The session began with a background interview (a), followed by a comprehensive tutorial on VR operations and AI probe usage (b). Subsequently, participants engaged in an AI probe study, initially focusing on planning a day's itinerary (c), then shifting to sharing stories about their campus life (d). The session culminated in a co-design session, encompassing both brainstorming activities (e) and thorough in-depth interviews (f).

3.4 Procedure

Our workshop included three parts as shown in Figure 2. The session lasted an average of 1.5 hours. The study was conducted in a laboratory setting. Two participants were in the same room, with two moderators present to assist in the experimental process. Before the start of the study, written informed consent was obtained from each participant, with the study purpose and task. All participants were informed of their right to withdraw from the study at any time, although no participants chose to do so. All participants received a payment of 100 RMB at the end of the study.

3.4.1 Part 1: Background Interview and Instruction. Our study began with an introduction of the study and tasks, along with a background information interview (e.g., age, communication frequency, VR experience, and AI experience). We then provided participants with the basic information and tasks about the study. To avoid potential frustration, we also informed them about the limitations of the AI probes, such as the possibility that AI responses might not fully meet their needs and that AI-generated suggestions might not always be correct.

In the instruction phase, we guided participants to put on the VR headsets, as shown in Figure 2-b. Considering that most participants had no prior experience with VR, we started by teaching them the basic operations of VR (e.g., selecting UI elements and manipulating virtual objects). Next, participants learned to use the AI features one by one. A virtual screen in front of participants displayed the description and instructions for each AI feature in text form in VR (as shown in Figure 1-6). Participants were required to follow these instructions to learn how to use the AI probe and click the “Learn Next” button to proceed to the next AI feature. During the learning session, two moderators observed and provided guidance as needed. Once participants were comfortable using the AI features, we moved them to the formal experimental tasks.

3.4.2 Part 2: AI Probe Study. After verifying participants' comprehension of the AI probes and VR manipulation, participants completed the following tasks: 1) Plan a day itinerary and 2) Share each other's campus life. These two tasks represented the two most common activities in inter-generational communication in everyday life: co-negotiation and life-sharing [29, 58]. This part was designed to be completed within an approximate timeframe of 30 minutes.

Task 1: Plan a day itinerary. The goal of this task was to stimulate discussion and decision-making between the two participants under a fixed topic. In the task, two participants collaborated to complete a day trip schedule (e.g., choosing and determining meals, attractions, transportation, etc.). They were provided with a blank schedule planner with a timeline marking several points in the day, which they were required to fill in by generating pictures or 3D objects, as shown in Figure 2-c. Participants were free to use any of the AI probes to complete the task.

Task 2: Share each other's campus life. This task encouraged participants to share and exchange stories about experiences unfamiliar to the other generation. Specifically, they took turns discussing their campus lives, with interruptions for clarification or engagement allowed to promote interactive dialogue. Participants did not have specific goals to accomplish and could stop the task at any time. They were also free to use AI probes to support their communication during the storytelling process.

3.4.3 Part 3: Co-design workshop. Finally, we conducted a co-design session with participants. First, the inter-generational pairs engaged in brainstorming sessions for about fifteen minutes without the presence of the moderators, as shown in Figure 2-e. They were required to discuss how the AI probes supported their communication in VR, the challenges they faced while using the AI probes, and their expectations for future interactions with AI in VR (i.e., preferred interaction strategies and the expected visual representations of AI). We provided paper, colorful sticky notes, markers, and various stickers to help them document their discussions.

After their discussion, two moderators entered the room to conduct in-depth interviews with both participants at the same time to gather detailed insights about the outcomes of their discussion. Our main objective was to encourage the participants to share the reasoning behind their conclusions.

3.5 Data Analysis

The qualitative analysis was conducted following the Grounded Theory Approach [15], aiming to generate a rich and empirical examination of how AI supports inter-generational communication in VR and how participants expect AI to be designed in the future. Following McDonald et al.'s guidelines for qualitative analysis in HCI [67], we prioritized identifying recurring concepts and categories over inter-rater reliability.

Specifically, our qualitative analysis contained the following steps: 1) Data Organization. Our data consisted of observational notes and audio recordings of the interviews. The observational notes were taken during the VR experiments, focusing on notable user behaviors and verbal expressions. After completing all experiments, we reviewed the experiment videos to supplement any missing details in the observational notes. The audio recordings were transcribed using DeepL⁶ and checked for accuracy by the primary authors; 2) Data Familiarization. All authors thoroughly read the observational notes and participants' narratives to understand common behaviors, perspectives, and expectations regarding AI use in VR; 3) Independent Coding and Category Formation. Each author independently created initial codebooks from three random transcripts [63]. Affinity Diagramming was then used to sort and integrate the coded information into preliminary categories [10]; 4) Category Refinement. Through iterative meetings, all authors discussed their codes and initial categories to identify discrepancies and make necessary revisions until a consensus was reached;

⁶<https://www.deepl.com>

5) Quote Extraction. One author extracted relevant quotes from all the data based on the refined categories; 6) Further Categorization. The authors further refined the categories and their relationships, using the extracted quotes to create a comprehensive description of how AI supports inter-generational communication in VR.

4 Findings

Table 2. Key Findings of RQ1: How AI-powered VR was leveraged to support inter-generational communication. The examples in the table were collected from observations in the study and participants’ envisioned applications of AI-powered VR in workshops. Envisioned scenarios were distinguished with an asterisk (*).

Key Findings	AI-powered VR Capabilities	Examples
AI-powered VR enhances mutual understanding	• Transforming abstract conversation content into concrete forms to better exchange ideas and thoughts	• Participants used Content Generation to generate an image of ‘chemistry laboratory’ to introduce their dream profession; used Info Assistant to explain ‘involution (nèi juǎn in Chinese)’
	• Providing various ways to interact with generated content to facilitate deeper communication	• Participants pointed to the roller in the AI-generated picture of the mimeograph machine and explained why the mimeograph machine could dirty the hands; interact with daily tools more intuitively for better imparting life skills to each other*
AI-powered VR fosters conversation fluency	• Providing interventions to avoid awkward silences, maintain balanced conversations, and manage potential friction	• Chat Facilitator suggested “ <i>Could you share something about your closest classmates?</i> ” when the conversation falls into silence
	• Personalizing AI representations to make its interventions more acceptable	• Participants recommended a realistic, context-appropriate AI appearance to provide convincing guidance for public activities; an inanimate virtual agent to reduce concerns about privacy and authenticity for private conversations; and friendly, cute virtual agents to create a relaxed and joyful atmosphere for casual talk*
AI-powered VR promotes active participation	• Providing content-specific spaces to offer an immersive and engaging conversation experience	• Content Generation switched the VR space to an old classroom when the OGs were reminiscing about their campus experiences
	• Modifying users’ avatars or virtual environments based on their emotions to ensure positive communication experiences	• Emotional Visualization could provide a bright and colorful environment in VR when the conversation is sad; offer vibrant and dynamic virtual effects in VR when the conversation becomes dull*

4.1 RQ1: How could AI-powered VR be leveraged to support inter-generational communication?

We found that AI-powered VR was leveraged to enhance mutual understanding, foster conversation fluency, and promote active participation. The key findings are summarized in [Table 2](#).

4.1.1 AI-powered VR enhances mutual understanding between two generations. Mutual understanding is important for inter-generational pairs to comprehend each other's thoughts, bridging generational knowledge gaps and enhancing empathy. We outlined two benefits of how AI-powered VR supports mutual understanding.

First, **AI-powered VR transforms abstract conversation content into concrete forms, such as images, videos, 3D models, and virtual scenes, enabling different generations to share and understand each other's ideas and thoughts more easily and intuitively.** During the probe study, participants used Content Generation to create images and 3D objects to introduce concepts that were unfamiliar to another generation. For example, when the older participants shared stories of their past school experiences, all of them tried to generate virtual scenes or objects, such as saying "Please take us back to my childhood classroom," (P4-OG) Similarly, when younger participants described their current jobs or future aspirations, they often generated images, such as "a chemistry laboratory" (P4-YG) or "a flag bearer raising a flag" (P1-YG). They indicated that incorporating visual aids would enhance clarity and prevent misunderstandings or confusion.

Additionally, participants found the Info Assistant useful to introduce era-specific vocabulary or events. This kind of vocabulary or events may be familiar to one generation but completely unfamiliar to the other. Participants suggested that AI played a crucial role in offering clear and neutral explanations, helping two generations to grasp the meanings and origins of these terms. For instance, P7-YG mentioned the term "involution (nèi juǎn in Chinese)" when sharing her college experience, which confused the older listener. To clarify, she employed the Info Assistant to offer a detailed background of the term.

Second, **AI-powered VR provides various ways to interact with generated content, such as pointing, grasping, zooming, and rotating, drawing listeners' attention to specific points and facilitating effective communication between generations.** P7-YG said, "In traditional phone calls, I find myself not fully engaged because I frequently struggle to understand what he is saying, making it easy for me to lose track of the conversation. ...But in this system, I stay engaged, constantly following the virtual content he shares. Additionally, by engaging with generated content, both parties experienced easier presentation and learning. When sharing campus life, P4-OG reminisced about using a mimeograph machine to produce exam papers, noting how it often left her hands dirty. She generated a picture of the "mimeograph machine", grasped the picture, and pointed to the roller, explaining to P4-YG: "You need to lay the paper flat on the template and then use an ink-filled roller to print it. You constantly have to adjust the template, which can get messy."

In the Co-design Workshop, participants envisioned that they could interact with AI-generated content in VR to impart life skills to each other remotely, enhancing inter-generational knowledge transfer. Compared to 2D screen-based communication, the 3D world of VR provides more realistic object displays and intuitive interactions that closely mimic real-life experiences, making it especially beneficial for older adults. Therefore, the older generations could effectively learn how to use the electronic tools from younger generations by generating these tools and practicing in VR. As the P6-OG indicated, "my daughter often brought me smart home appliances like robotic vacuum cleaners to ease my workload. However, we struggled to use them and often abandoned them when she was not home." They found voice or video calls inefficient for troubleshooting, but VR-generated 3D objects could help bridge this gap, allowing them to practice without real-world consequences. Similarly, younger participants also indicated their need to seek advice from their elders on life skills, such

as “*learning grandma’s signature dish*” (P6-YG). They believed that VR Content Generation could effectively aid them in learning life skills hands-on from their elders.

4.1.2 AI-powered VR fosters conversation fluency between two generations. We identified two ways in which AI-powered VR facilitates conversation fluency.

AI-powered VR could suggest topics or provide timely interventions based on contextual content, helping two generations avoid awkward silences, maintain a balanced conversation, and manage potential friction. AI could comprehend user dialogues and **provide potential topics and activities** to avoid awkward silences and keep the conversation flowing. During the study, eight pairs of participants encountered silence during their interactions. For example, after P1-YG stated he was doing well in school, both he and the OG participant fell silent. Sensing this, the Chat Facilitator intervened with a voice suggestion, “*Could you share something about your closest classmates?*” Participants appreciated this mediator and felt it kept the conversation active. Participants envisioned that AI could suggest mutually interesting topics based on their personalities, backgrounds, and previous conversations.

Participants perceived that AI could **balance the participation** of two generations by providing real-time interventions. During the study, AI was activated in seven pairs of participants to balance their conversation. For example, in Task 1, younger participants may independently craft travel plans without seeking advice from their older participants. In this instance, the Chat Facilitator provided suggestions using a voice prompt: “*Let’s have both sides discuss together the places you would like to visit.*” They appreciated the reminder of AI and thought it made them attentive to the other’s feelings during the conversation. Conversely, in Task 2, the older participants often engaged in extended storytelling, ignoring whether the younger listeners were interested. Although the younger participants generally refrained from interrupting the storytelling, believing it to be impolite and disrespectful towards the older generations, they hoped that AI could serve as a mediator to ensure the harmonious progression of the conversation. The older participants also indicated their welcoming of AI’s reminder. As P5-OG described, “*I probably didn’t realize this problem because I was excited to tell my story at this time. I’m grateful for the AI’s reminder. I think discussing with him together is much more enjoyable than just talking by myself.*”

Participants also envisioned that AI could **offer fair interventions to manage potential friction**, maintaining a harmonious atmosphere in communications. During the co-design workshop, participants suggested using AI as a mediator in daily conversations to mitigate conflicts. In this case, AI could intervene in conversations by suggesting a temporary cool-down or offering objective solutions. For instance, in task 2, P8-OG opined that today’s campus life is happier than before, but P8-YG countered, citing increased academic pressure now. This led P8-OG to criticize P8-YG for lacking resilience and endurance, causing unpleasantness in their conversation. P8-YG suggested that “*this was a moment where AI could intervene.*” Intriguingly, several participants expected they might prefer AI over human coordinators for resolving inter-generational conflicts. This preference stemmed from their perception of AI as being neutral and unbiased, unlike humans who could be influenced by emotional attachments or social connections. For instance, P8-YG remarked, “*When I have conflicts with my grandfather, I prefer listening to AI over my parents or grandmother, as they tend to take sides and can’t mediate impartially.*” This highlighted the perceived impartiality of AI in conflict resolution.

AI-powered VR provides AI with embodied and personalized avatars that adapt to different communication contexts, making AI’s intervention more acceptable for users. In the workshop, participants noted that while they found verbal interventions helpful for facilitating communication, a visual representation of AI would increase their trust, as it seemed more human-like and reliable. P5-OG emphasized, “*In Task 1, without an embodied avatar, the AI’s suggestions*

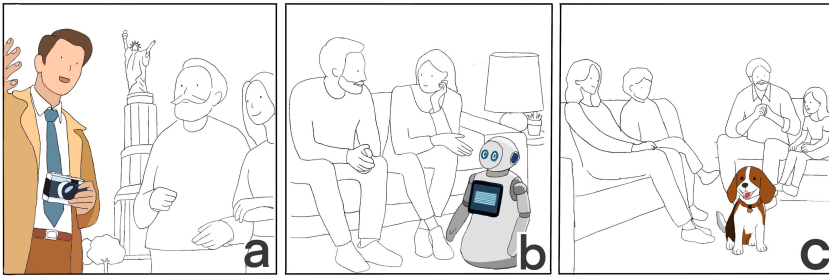


Fig. 3. Illustration of participants' suggestions for AI's visual representation: a) shows a realistic, context-appropriate AI appearance such as a tour guide describing the attraction. Participants indicated that such representations would be especially beneficial in VR-shared activities, enhancing immersion and providing convincing guidance; b) depicts an inanimate virtual agent as a robot. Participants believed that it reduces their concerns about privacy and authenticity, making them more suitable for private conversations; c) represents a friendly and cute pet as a virtual agent. Participants believed that such representations were effective in easing tension and adding humor to interactions.

felt as if they suddenly emerged from the air, which seemed abrupt and made us hesitate in accepting the suggestions. The embodied avatar allowed users to feel they were interacting with a real person rather than an intangible system. Moreover, avatar appearances can be customized according to user needs and interaction scenarios, enabling AI to integrate seamlessly into conversations and adapt dynamically. Participants proposed three distinct types of visual representations, each suited to different scenarios, as shown in Figure 3.

First, participants suggested a **realistic, context-appropriate AI appearance**, such as a virtual tour guide or a teacher. They believed such representations would enhance immersion and provide convincing guidance in VR-shared activities. As P4-YG noted: *"A guide dressed in formal attire in a VR museum is more convincing than a dog or cat."* However, this realistic avatar may hinder casual inter-generational conversation due to its strong and invasive presence, as P6-YG commented: *"Such a strong presence of AI feels like it's listening in on everything we say, which is somewhat uncomfortable."*

The second proposed form was an **inanimate virtual agent**, such as a robot. Participants indicated that, in contrast to lifelike avatars, the non-human appearance of robots reduced their concerns about privacy and authenticity, making them more suitable for private conversations. Additionally, robots were seen as impartial, free from emotions or biases, and thus, ideal for conflict mediation. As P3-YG highlighted *"I feel more open to listening to the unbiased advice from the emotionless robots."*

Lastly, participants suggested **friendly and cute virtual agents**, like cartoon characters or pets, for their ability to create a relaxed and joyful atmosphere. This approach may mediate conversational dynamics, especially during awkward or silent moments between generations. These avatars were seen as effective in easing tension and adding humor to interactions. As P1-YG remarked, *"I prefer cartoon puppies as they feel like part of the family and their interactions are natural and unforced."*

4.1.3 AI-powered VR promotes active participation between two generations. We identified two reasons why AI-powered VR made inter-generational communication more active.

AI-powered VR provided content-specific immersive environments and engaged both participants in active conversation. Participants commended the benefits of switching the virtual scenes in VR according to their communication content. They thought that it provided an immersive communication experience that fostered a greater willingness to listen and share

stories. Participants indicated that their current storytelling process was usually led by the older generations while the younger generations often lacked a sense of involvement and empathy. As P10-YG mentioned *“Sometimes, I get impatient listening to her youth stories because they’re hard for me to relate to... However, I found the stories more engaging during VR experiences, as they make me feel like I’m reliving those moments with her.”* They appreciated that AI could generate scenes in VR (e.g., an old family home or a meaningful location from the past) or 3D virtual items (e.g., agricultural tools or stroller) to enhance the storytelling experience and actively engage listeners.

Additionally, participants envisioned that AI could provide active co-learning experiences by creating historical scenes and geographical landscapes mentioned in textbooks within VR environments. They indicated that co-learning was a common activity they engaged in at home, and AI could potentially benefit this activity by boosting learning interests and outcomes. For example, the P12-OG expressed that *“AI can help bring to life the scenes described in ancient poetry, allowing us to immerse ourselves in these landscapes witnessed by poets centuries ago in VR.”*

AI-powered VR can modify the users’ avatars or virtual environments based on their emotions to ensure positive communication experiences. During the VR experience, participants’ pronounced emotions were visualized in real-time.

Participants appreciated the visualization of their positive emotions. P7-YG participant expressed, *“It made me feel my uncle’s joy, which in turn made me happier and more willing to discuss the topic further with him”*. Participants suggested two ways AI could help express their happiness. The first way is to automatically modify the expression of users’ avatars, such as *“showing hearty eyes, wide smiles, or even playful winks”*. Embodied avatars, as virtual representations of users, can effectively influence the social experience during interactions [103]. Participants believed that this subtle intervention could create a pleasant communication atmosphere without disrupting the flow of the conversation. The second approach is to modify VR environments to exaggerate users’ positive emotions, such as *“providing joyful background music, altering the surrounding environment, and making objects around them dance”*. Participants felt that this visual enhancement of positive emotions could effectively increase their emotional involvement with each other. They believed that immersive VR could make these enhancements more impactful, providing a more realistic experience than screen-based communication platforms.

While most participants saw value in visualizing positive emotions to enhance the conversation atmosphere, they unanimously preferred not to exaggerate or display their negative emotions to others. As P9-OG mentioned, *“Imagine feeling down and suddenly there’s a storm in the VR — that’d just make us feel worse.”* They suggested that if the conversation is serious or sad, the scene could transform into something bright and colorful, like a sunny park or a lively beach, to lift spirits and re-energize the dialogue.

4.2 RQ2: What are the challenges of using AI to support VR-based inter-generational communication?

We identified three challenges participants faced when using AI to support their communication in VR, including the inappropriate timing and manner of AI interventions in VR, the difficulties of AI in comprehending participants’ intentions, and participants’ over-reliance on AI.

4.2.1 Participants felt that the inappropriate timing and manner of AI interventions in VR disrupted the conversation flow and disturbed the harmonious conversational atmosphere. Participants perceived that VR’s immersive environment makes public interventions more prominent and effective in highlighting issues. Consequently, AI’s public interventions in highlighting others’ problems become more inappropriate and impolite, particularly when involving the elderly. As P9-YG expressed, *“We need to be respectful when talking with elders. Even though AI’s public intervention*

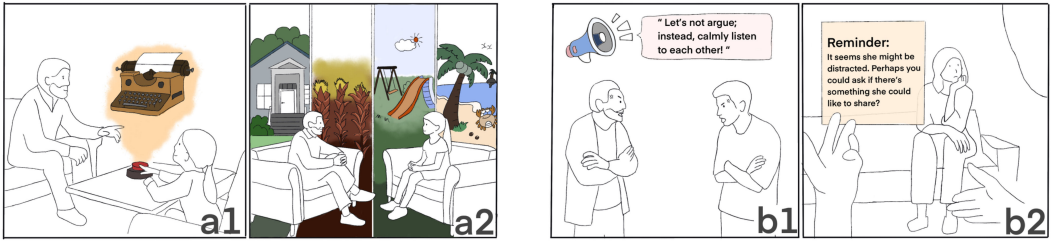


Fig. 4. Illustrations of participant-proposed AI interaction strategies in VR. Two activation methods were envisioned on the left: a1) passive activation, which emphasizes user autonomy and aims to minimize interruptions; and a2) proactive activation, where the virtual environment in VR dynamically changes in response to the user's emotions or conversation content. Two intervention strategies are proposed on the right: b1) prompts visible to both parties involved in the communication; b2) private prompts designed to discreetly remind the user of their behavior.

stopped my grandfather from going on and on, it made the conversation a bit awkward afterward. Additionally, participants reported that AI sometimes intervened at inappropriate times during their conversations, disrupting the natural flow and causing discomfort. P12-YG indicated, *"While my grandparent sharing his school story, the AI suddenly interrupted. We were all engaged and didn't need any assistance."*

Participants suggested that future designers could make AI interventions more appropriate and acceptable in VR. They envisioned two AI activation approaches and two AI intervention forms, as shown in Figure 4. They also outlined how these could be applied in inter-generational interactions.

Generally, participants favored a manual, user-controlled **passive method** to invoke AI's features for most scenarios (Figure 4-a1), which enables users to engage the AI as needed through buttons, voice commands, or gestures. Such control is essential for maintaining user autonomy and minimizing interruptions, especially during private or in-depth conversations. Nevertheless, participants saw the potential for a **proactive method** for specific situations like emotional downturns (Figure 4-a2). They noted that during intense emotions such as sadness or anger, they tended to become absorbed in their feelings, focusing on the emotions rather than seeking ways to alleviate or manage them. In these situations, participants appreciated the idea of proactive AI intervention. P11-OG stated that: *"When I become angry or sad during a conversation, I hope VR can display beautiful scenery to help regulate our emotions and improve the relationship."*

In terms of intervention manners, participants believed **publicly displayed** intervention could be beneficial for providing objective solutions and information to facilitate communication, as they can help clarify misunderstandings and ensure that both parties are on the same page. For example, if a misunderstanding leads to an argument, AI could intervene with prompts noticeable to both parties, such as an audio announcement to suggest a cool-down period or provide an objective solution. Additionally, participants suggested that the **private prompts** is more acceptable and polite to remind them in situations where one party dominates the conversation or says something that makes the other person feel uncomfortable (Figure 4-b2), especially for the elderly. For instance, a discreet text prompt such as *"You've spoken at length; let the other person share their views"* could be displayed only in the dominant speaker's VR environment; Similarly, when one party discusses uninteresting topics for the other, a prompt like *"It seems that they don't want to talk about this. Try asking about their enjoyable moments at school."* could be shown to the speaker. Participants indicated that the private prompts, compared to cues visible to both parties, would maintain a positive conversation atmosphere and flow while preserving the dignity of the prompted user.

Older participants also saw value in private prompts, acknowledging that they sometimes did not realize the younger generation's disinterest in their topics. P9-OG stated, *"If I'm speaking excitedly, I might not realize that others aren't getting a chance to speak. I hope the AI can remind me discreetly."*

4.2.2 Participants found that AI sometimes misunderstands their intentions and provides wrong responses, impacting inter-generational communication efficiency and potentially heightening conflicts. We attributed this challenge to two factors:

Firstly, **the lack of non-verbal cues** led to failures in understanding participants' intentions in VR. Compared to screen-based communication tools, users are more likely to express non-verbal cues — such as gestures, facial expressions, head orientation, gaze direction, and proximity — when communicating with others in VR [1, 92]. In the study, we observed that participants unconsciously employed non-verbal cues during their interactions with the AI. For example, P4-OG pointed at a picture in VR and asked the AI, *"Which section of the Great Wall is this?"* The AI was unable to answer the question because it did not recognize the participants' pointing gestures and pointing areas. Participants expressed a desire to integrate multi-modal user inputs in VR, such as body language, facial expressions, gestures, gaze, and interactions, to better perceive their intention when communicating with AI.

Secondly, AI's **limited understanding of participants' background**, including familial dynamics, communication habits, and regional language, led to conversation disruptions and may even exacerbate conflicts. Participants acknowledged the complexities inherent in each family, posing a substantial challenge for AI in accurately understanding conflicts and responding suitably. Consequently, they worried that AI might intervene inappropriately in situations where it should not. For example, P9-YG mentioned, *"I think it is important for AI to understand that my grandfather is accustomed to speaking loudly due to hearing loss, which should not be interpreted as a sign of anger."* Additionally, participants voiced concern that AI might make erroneous interventions because it may not fully grasp the dynamics of conflicts. As P11-YG mentioned, *"Sometimes, a simple remark from my dad can make me angry. This is because there are long-standing sibling rivalries or past incidents that substantially impact the dispute. I worry that AI, lacking family background, might not understand the root of my anger and provide effective advice or solutions."* Participants also highlighted the AI's inability to recognize dialects impairing their communication experience. This issue was particularly evident in China, where many older adults, especially those residing in rural areas, struggled with fluency in standard Mandarin. Our study revealed a strong preference among most participants for using local dialects when communicating with family members. However, to facilitate better understanding by the AI during the study, they were compelled to switch from their native dialects to standard Mandarin, which disrupted the natural flow of conversation.

4.2.3 Participants' over-reliance and perceived fairness in AI led to negative consequences. During the experiment, we found that participants sometimes over-trusted the results provided by AI, leading to unreasonable decisions. For instance, when P6 requested a one-day tour plan for Beijing, the AI provided an extremely packed and overwhelming schedule. The suggested itinerary included visiting the Forbidden City, Tiananmen Square, and the Great Wall in the morning, followed by the Summer Palace, Beihai Park, and the Beijing Film Museum in the afternoon. Despite the impracticality, three groups of participants accepted the AI's suggestions and incorporated these activities into their plan.

Furthermore, during the interviews, two groups of participants mentioned that they perceived AI as fair and unbiased, unlike humans who could be influenced by emotional attachments or social connections. Consequently, they were more inclined to trust AI interventions over those from human facilitators. However, previous research indicates that AI-generated results can contain

social biases [60, 85]. Without critical judgment of the AI's outputs, this could lead to negative consequences, such as reinforcing existing biases or making flawed decisions.

5 Discussion

In this section, we further discuss our findings along the following aspects: 1) the advantages of leveraging AI in VR to enhance inter-generational communication and how this integration addressed issues of inter-generational communication, 2) the challenges of integrating AI to support inter-generational communication in VR and associated design implications, and 3) a comparison of AI facilitators with human facilitators to understand their respective strengths, weaknesses, and suitability in different contexts.

5.1 The integration of AI and VR to support inter-generational communication

This work aimed to understand how inter-generational pairs perceive and leverage AI to support communication in VR (RQ1). Previous works have demonstrated the advantages of VR in facilitating inter-generational communication, such as achieving a degree of co-presence and companionship by immersive shared environment [2, 102], cultivating empathy by embodying avatars with each other's appearance [79, 91], or providing a casual and relaxed atmosphere by engaging in various activities [36, 102]. This work extends prior work by leveraging AI to create a *content-aware VR environment* to support a dynamic communication process. Through a probe-based participatory design workshop, we identified three key ways in which inter-generational users wish to leverage AI in VR to support communication.

First, participants recognized the potential of the integration of AI and VR to *enhance mutual understanding* by bridging knowledge gaps between generations. Prior research highlights a widespread desire to maintain familial connections and share life experiences with family members [73, 75, 97]. However, the diverse cultural backgrounds, characterized by unique values, beliefs, and communication styles, often lead to misunderstandings between generations [73, 97]. Our findings suggest that AI and VR could facilitate *mutual understanding* by allowing users to switch abstract conversation content into concrete objects to effectively exchange their ideas. Additionally, users could further interact with the AI-generated outcomes to discuss and clarify their thoughts. Such interaction could not only address the knowledge gaps but also enrich the life-sharing and storytelling experience, making it more vivid and tangible.

Second, the integration of AI and VR was seen as pivotal in *enhancing conversation fluency* of two generations through appropriate and real-time interventions. Typically, non-fluent conversations between two generations could result in brief and unsatisfying interactions, limiting opportunities for meaningful knowledge and cultural exchange [58, 102]. Our research demonstrates AI's effectiveness in avoiding awkward silences, maintaining balanced conversations, and alleviating potential conflicts. When integrated with VR environments, AI can embody various visual representations (e.g., teacher, robot, puppy) based on the conversation context, making it easier for users to accept AI's suggestions.

Third, participants suggested that the integration of AI and VR could *promote active participation* between generations. Prior research suggests that the lack of shared knowledge and the subtlety of emotional expression hinder active engagement across generations, leading to miscommunication, a lack of genuine connection, and a sense of impatience among the youth [32, 48, 80, 104]. Participants noted that AI could modify the VR environment in real-time based on the context content and emotional tone of conversations, encouraging users to share and listen more attentively to each other, thus enriching the quality of the interaction overall.

Overall, our findings indicate that integrating AI with VR opens new avenues for enhancing communication between generations, providing a promising direction for future studies that aim to use VR in supporting inter-generational communication.

5.2 Design implications for AI to better support inter-generational communication in VR

This work revealed the challenges of using AI in VR to support inter-generational communication (RQ2), including inappropriate timing and manner of AI interventions in VR, AI's inaccurate understanding of user intentions, and the issues arising from over-reliance and trust in AI. Based on these user-identified challenges, we outlined key design implications to guide future researchers and developers in creating AI-based VR platforms that enhance inter-generational communication. Among these design implications (DIs), DI1-4 focuses on how to effectively combine the advantages of AI and VR to enhance inter-generational communication experiences. Meanwhile, DI5-7 are key recommendations derived from our study for using AI to support inter-generational communication, which may apply to VR and other communication platforms.

5.2.1 Exploring various VR designs for providing appropriate AI interventions to support inter-generational communication. While participants appreciated the benefits of integrating AI and VR, they also highlighted issues with AI's inappropriate intervention. Based on participant feedback and suggestions, we identify three promising directions for future work.

First, we suggest that **DI1: AI could adapt different representations in VR based on conversational contexts, thereby creating a suitable communication atmosphere and making their suggestions more acceptable.** Similar to users' avatars that can influence others' perceptions [50, 51, 68], in our study, participants indicated that AI's visual representations in VR could also affect their perceptions of AI's interventions. They recommended three kinds of AI representations for different communication contexts: a realistic, context-appropriate AI appearance to provide convincing guidance for public activities; an inanimate virtual agent to reduce concerns about privacy and authenticity for private conversations; and friendly, cute virtual agents to create a relaxed and joyful atmosphere for casual talk. This aligns with HCI research on multi-agents, emphasizing personalized interaction and adaptability in social contexts [14, 40].

Second, we suggest that **DI2: AI's intervention could be enriched using various VR visual effects, creating engaging and appealing social experiences.** Prior studies have demonstrated the effectiveness of diverse visual representations — such as emotes, particles, creatures, fur, skeuomorphic objects, ambient light, and halos — in conveying users' social behaviors and emotional states in VR [9, 22, 52, 87]. They show that diverse VR visual effects could make interactions more active and fun, potentially leading to positive emotional states and a greater willingness to communicate [72, 86, 103]. These novel approaches offer exciting possibilities for future research.

Finally, we suggest that **DI3: future research should leverage appropriate forms of AI intervention in VR (i.e., public or private prompts, and automatic or manual activation) to maintain the smoothness of communication.** In our study, participants found that the public interventions of AI sometimes disrupted the flow of communication and could potentially lead to a negative atmosphere. They thought that prompts visible to both parties could be beneficial for providing objective solutions and information. On the other hand, private prompts are suitable for discreetly reminding one user during awkward silences or imbalances in the conversation, thus preventing potential embarrassment or discomfort. Selecting the appropriate form of AI intervention in VR helps maintain a positive conversation atmosphere and flow, ensures that interactions remain engaging, and preserves the dignity and comfort of the prompted user.

5.2.2 Enhancing AI's interpretation of user needs in VR. As mentioned in Section 4.2, participants often interacted with the AI using non-verbal cues in VR. However, AI failed to understand the users' non-verbal cues correctly, thereby providing incorrect responses. This insight led us to propose **DI4: AI needs to incorporate users' non-verbal cues, such as postures, gestures, orientations, finger-pointing, and gaze direction, to understand user intentions precisely.** VR can easily track multiple channels of user non-verbal information in VR since users wear headsets and hold controllers [62, 95, 99]. By incorporating these non-verbal cues, AI can better understand user intentions and provide more accurate responses.

Additionally, we propose **DI5: AI needs to understand the communication context and dynamics among users to better interpret their needs and provide appropriate responses.** Communication within families is often more casual and centered around personal topics, a relaxed tone, and language reflecting familial backgrounds [34, 58]. However, participants reported difficulties with AI comprehending this informal style. They voiced concerns that the AI may misunderstand family-specific expressions or humor, resulting in confusion and disrupting the flow of conversation. Therefore, it is important for AI to grasp the context and subtleties of users' conversations. To achieve this, future AI designers should ask users to provide basic information during the initial setup phase, such as educational backgrounds, cultures, and common languages. Additionally, for ongoing effectiveness and respectfulness, AI should continuously learn and update its knowledge and algorithms to stay in sync with new information, evolving social norms, and family-specific events [7, 107, 112]. This detailed understanding would enable AI to grasp the context and nuances of users' speech more accurately. This not only helps AI provide personalized responses that resonate with each user's unique experiences and viewpoints but also helps AI determine appropriate timing for intervention, thereby avoiding interruptions to the natural flow of inter-generational communication.

5.2.3 Ensuring trustworthy and unbiased AI in inter-generational communication. In this study, we observed that some participants placed substantial trust in the information provided by AI, viewing it as a fair and unbiased facilitator. However, based on previous research, we recognize that such trust may lead to certain issues for two primary reasons. Firstly, the AI's outputs are not always accurate [47]. Blind acceptance of AI-generated results can mislead conversations and lead to erroneous decisions [28, 47]. Secondly, AI is not unbiased [60, 85]. LLMs used in AI are trained on extensive datasets that often contain inherent social biases [19, 20, 78]. These biases can inadvertently influence the AI's behavior and outputs, leading to unintended consequences. For instance, an AI designed to facilitate conversation might unintentionally favor certain cultural norms or gender stereotypes embedded in its training data, thus affecting the quality and fairness of its interactions [74]. Therefore, if users unconditionally trust AI's results, not only can it lead to inefficient communication but also reinforce stereotypes or create friction in conversations, particularly when different generations have varying levels of digital literacy and cultural perspectives.

We propose two design considerations for future AI designers to mitigate these issues. First, **users should be reminded to approach AI-generated results with a critical mindset during their interactions (DI6).** Several strategies could be employed to achieve this. Initially, designers could set clear expectations that AI-provided answers may not always be accurate before users engage with the system [85]. Additionally, interactive tutorials can be used to educate users on the AI's strengths and limitations, including examples of accurate and inaccurate responses. Throughout the interaction, continuous reminders of the AI's limitations can be provided through periodic notifications or prompts [3]. For example, after delivering a response, the AI could include a brief message suggesting that users verify the information independently or consider alternative viewpoints. Another approach is to incorporate a feature that highlights the AI's confidence level in

its responses [12]. This could be represented as a confidence score or a visual gauge that indicates how certain the AI is about the information provided.

Second, **it is essential to implement bias detection and correction mechanisms within the AI system (DI7)**. Previous studies have demonstrated potential ways of identifying and mitigating biases within AI systems. For instance, regular audits of the AI's outputs, inclusive training datasets, and user feedback loops are vital for ensuring a more balanced and fair interaction environment [57, 70]. Additionally, periodic reviews can identify and rectify biases in the AI's responses, ensuring they remain accurate and impartial. Using diverse and representative training datasets can help mitigate embedded social biases [57, 90, 108], while real-time user feedback can provide immediate insights into potential issues, allowing for swift adjustments. Mechanisms such as feedback buttons or surveys can empower users to report biased or inappropriate responses, ensuring ongoing refinement of the AI [88]. Addressing these biases is important for the development of practical and meaningful AI-based VR platforms that genuinely enhance inter-generational communication without perpetuating existing societal biases.

5.3 Comparing AI+VR with co-located human facilitators in remote inter-generational communication

Previous research indicates that remote inter-generational communication sometimes relies on co-located human facilitators to coordinate conversation [30, 35, 41, 82, 109]. These human facilitators, who might be parents, siblings, or grandparents, frequently manage three-party communication via phone or video calls to facilitate inter-generational exchanges. Such facilitators are often necessary in situations where, for example, a grandchild may not be very familiar with their grandparent living apart, leading to potential awkwardness and a lack of topics to discuss if they converse directly. In these scenarios, parents often act as facilitators to join conversations with their children and grandparents. They offer assistance including providing context [30], bridging generational understanding gaps [41], engaging both generations' attention [30, 82], and suggesting topics to initiate conversation [41, 109]. These studies show that human facilitators help maintain effective inter-generational communication, enhance mutual understanding between generations, and strengthen family bonds. However, they also highlight several drawbacks associated with human facilitators. For instance, facilitators may sometimes dominate conversations, steering them toward unwanted topics [41, 109]. Additionally, inter-generational pairs may find it uncomfortable to discuss specific topics or personal matters in the presence of the facilitators [30, 41].

Our findings suggest that AI-empowered VR can effectively complement or enhance the coordination offered by human facilitators, such as facilitating mutual understanding, enhancing engagement, and initiating discussions. Furthermore, AI facilitators in VR can mitigate some drawbacks associated with human facilitators [30, 41, 109]. For example, by adjusting the level of AI intervention, AI facilitators could avoid dominating conversations, ensuring a comfortable interaction experience. Additionally, the AI can adopt different virtual representations to minimize its presence to help users feel more at ease when discussing private matters. Research also indicates that VR's immersive environment and embodied avatars offer a greater sense of companionship and togetherness, alleviating feelings of longing between generations [102].

However, we also found that AI has shortcomings compared to human facilitators. Due to a limited understanding of users' personalities and family histories, AI might fail to detect issues in conversations accurately, potentially disrupting conversations [64, 109]. Additionally, AI may lack flexibility when handling unexpected situations, especially those beyond its training scope [108]. Human facilitators, on the other hand, exhibit high flexibility and adaptability, improvising in response to sudden changes and complex scenarios. Moreover, AI facilitators may raise privacy concerns [65, 112]. Users might be wary of how their data is used, affecting trust. Human facilitators can

build trust through direct interaction and personal connection, providing clearer privacy assurances and reducing concerns about data security. Furthermore, AI facilitators cannot update each other's status [30, 41] or coordinate varied family schedules as effectively as human facilitators [35, 109]. By addressing these areas, future research can contribute to optimizing the roles of AI facilitators in enhancing inter-generational communication, ensuring that they complement and enhance the benefits provided by human facilitators while minimizing potential drawbacks.

6 Limitations and Future Work

We acknowledge five limitations that motivate potential future research directions:

First, our study was conducted with inter-generational pairs from Chinese communities, which might introduce cultural and regional biases to our results [106, 110]. For instance, the optimal ways for AI intervention in inter-generational conversations might vary across different countries. While our findings highlight general AI perceptions and implications in this context, further research with a diverse international sample is crucial to understanding the varying impacts of cultural differences on these perceptions.

Second, this work mainly focused on understanding users' perceptions and needs of AI in enhancing inter-generational communication. We did not conduct comparative evaluations to gauge AI's effectiveness against a standard task completion baseline. Building on our findings, future research should explore the impact of AI-powered VR applications more thoroughly. This could involve using control groups in subsequent studies for a clearer understanding of how AI interventions specifically aid inter-generational communication.

Third, to prevent long rendering times from disrupting user communication, we employed pre-generated content instead of real-time 3D models and virtual scenes in our study. While this method allowed for controlled and consistent experimental conditions, it limited our ability to explore certain characteristics of AI such as unpredictability and latency in the context of inter-generational communication. Future research should explore the challenges and impacts of using real-time AI generation in VR to gain a more comprehensive understanding.

Fourth, to ensure consistent experimental conditions and minimize distractions related to avatar appearance, we did not employ an avatar customization function in our study. While this decision facilitated a more focused interaction with AI-powered VR, it restricted our investigation into how discrepancies between user and avatar appearance could affect the sense of embodiment and overall immersive experience. Future research should consider incorporating customizable avatars to examine their influence on user engagement and the dynamics of inter-generational communication.

Moreover, although our AI probe-based co-design workshop revealed interesting insights, it lasted for a relatively short period and was conducted in a controlled laboratory setting. These factors may compromise the authenticity and practical applicability of the experimental outcomes [71]. Future research should focus on longer, more extensive, and ecologically valid studies outside the lab to better understand AI's long-term effects and gather practical insights for daily use in inter-generational communication.

7 Conclusion

Inter-generational communication is crucial for enhancing family bonds. Through a probe-based participatory design study with twelve inter-generational pairs, our work examines how AI can enhance inter-generational communication in VR by improving mutual understanding, conversation flow, and participation. It also highlights challenges, such as inappropriate AI interventions, misinterpretation of user intentions, and over-reliance on AI. We provide design implications to

address these issues, aiming to develop AI tools that are both technologically advanced and sensitive to emotional and family dynamics, fostering stronger family connections.

Acknowledgements

This work is partially supported by the National Key R&D Program of China(No. 2022YFF0902303), the 2022 major science and technology project “Yuelu Multimodal Graph-Text-Sound-Semantic Gesture Big Model Research and Demonstration Application” in Changsha (No. kh2301019), the Guangzhou-HKUST(GZ) Joint Funding Project (No. 2024A03J0617), Guangzhou Science and Technology Program City-University Joint Funding Project (No. 2023A03J0001), the Project of DEGP (No.2023KCXTD042), and Guangdong Provincial Key Lab of Integrated Communication, Sensing and Computation for Ubiquitous Internet of Things (No. 2023B1212010007).

References

- [1] Ahsan Abdullah, Jan Kolkmeier, Vivian Lo, and Michael Neff. 2021. Videoconference and Embodied VR: Communication Patterns across Task and Medium. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–29.
- [2] Tamara Afifi, Nancy Collins, Kyle Rand, Chris Otmar, Allison Mazur, Norah E Dunbar, Ken Fujiwara, Kathryn Harrison, and Rebecca Logsdon. 2022. Using Virtual Reality to Improve the Quality of Life of Older Adults with Cognitive Impairments and their Family Members who Live at a Distance. *Health Communication* (2022), 1–12.
- [3] Abhishek Aggarwal, Cheuk Chi Tam, Dezhi Wu, Xiaoming Li, and Shan Qiao. 2023. Artificial intelligence-based chatbots for promoting health behavioral changes: Systematic review. *Journal of medical Internet research* 25 (2023), e40789.
- [4] Khair Al-Kodmany. 2001. Bridging the gap between technical and local knowledge: Tools for promoting community-based planning and design. *Journal of Architectural and Planning research* (2001), 110–130.
- [5] 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 (2021), 1–25.
- [6] Holly B Ballenger. 2014. Photography: a communication tool. (2014).
- [7] Gagan Bansal, Besmira Nushi, Ece Kamar, Walter S Lasecki, Daniel S Weld, and Eric Horvitz. 2019. Beyond accuracy: The role of mental models in human-AI team performance. In *Proceedings of the AAAI conference on human computation and crowdsourcing*, Vol. 7. 2–11.
- [8] Ivo Benke, Michael Thomas Knierim, and Alexander Maedche. 2020. Chatbot-based emotion management for distributed teams: A participatory design study. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (2020), 1–30.
- [9] Guillermo Bernal and Pattie Maes. 2017. Emotional beasts: visually expressing emotions through avatars in VR. In *Proceedings of the 2017 CHI conference extended abstracts on human factors in computing systems*. 2395–2402.
- [10] Hugh Beyer and Karen Holtzblatt. 1999. Contextual design. *interactions* 6, 1 (1999), 32–42.
- [11] Mariana J Brussoni and Susan D Boon. 1998. Grandparental impact in young adults’ relationships with their closest grandparents: The role of relationship strength and emotional closeness. *The International Journal of Aging and Human Development* 46, 4 (1998), 267–286.
- [12] Tommy Bruzzese, Irena Gao, Griffin Dietz, Christina Ding, and Alyssa Romanos. 2020. Effect of confidence indicators on trust in AI-generated profiles. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–8.
- [13] Chris Bussell, Ahmed Ehab, Daniel Hartle-Ryan, and Timo Kapsalis. 2023. Generative AI for Immersive Experiences: Integrating Text-to-Image Models in VR-Mediated Co-design Workflows. In *International Conference on Human-Computer Interaction*. Springer, 380–388.
- [14] Rafael C Cardoso and Angelo Ferrando. 2021. A review of agent-based programming for multi-agent systems. *Computers* 10, 2 (2021), 16.
- [15] Kathy Charmaz. 2006. *Constructing grounded theory: A practical guide through qualitative analysis*. sage.
- [16] Dana Cohen-Bar, Elad Richardson, Gal Metzer, Raja Giryes, and Daniel Cohen-Or. 2023. Set-the-Scene: Global-Local Training for Generating Controllable NeRF Scenes. *arXiv preprint arXiv:2303.13450* (2023).
- [17] Raymundo Cornejo, Mónica Tentori, and Jesús Favela. 2013. Enriching in-person encounters through social media: A study on family connectedness for the elderly. *International Journal of Human-Computer Studies* 71, 9 (2013), 889–899.
- [18] Natalie Cotterell, Tine Buffel, and Christopher Phillips. 2018. Preventing social isolation in older people. *Maturitas* 113 (2018), 80–84.

- [19] Roxana Daneshjou, Mary P Smith, Mary D Sun, Veronica Rotemberg, and James Zou. 2021. Lack of transparency and potential bias in artificial intelligence data sets and algorithms: a scoping review. *JAMA dermatology* 157, 11 (2021), 1362–1369.
- [20] Taher Dehkharghanian, Azam Asilian Bidgoli, Abtin Riasatian, Pooria Mazaheri, Clinton JV Campbell, Liron Pantanowitz, HR Tizhoosh, and Shahryar Rahnamayan. 2023. Biased data, biased AI: deep networks predict the acquisition site of TCGA images. *Diagnostic pathology* 18, 1 (2023), 67.
- [21] Joseph A DeVito and Joe DeVito. 2007. The interpersonal communication book. (2007).
- [22] Arindam Dey, Thammathip Piumsomboon, Youngho Lee, and Mark Billinghurst. 2017. Effects of sharing physiological states of players in collaborative virtual reality gameplay. In *Proceedings of the 2017 CHI conference on human factors in computing systems*. 4045–4056.
- [23] Paola Di Giuseppantonio Di Franco, Carlo Camporesi, Fabrizio Galeazzi, and Marcelo Kallmann. 2015. 3D printing and immersive visualization for improved perception of ancient artifacts. *Presence: Teleoperators and Virtual Environments* 24, 3 (2015), 243–264.
- [24] Wen Duan, Naomi Yamashita, Yoshinari Shirai, and Susan R Fussell. 2021. Bridging fluency disparity between native and nonnative speakers in multilingual multiparty collaboration using a clarification agent. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–31.
- [25] Guido A Entenberg, Sophie Mizrahi, Hilary Walker, Shirin Aghakhani, Karin Mostovoy, Nicole Carre, Zendrea Marshall, Gilly Dosovitsky, Daniellee Benfica, Alexandra Rousseau, et al. 2023. AI-based chatbot micro-intervention for parents: Meaningful engagement, learning, and efficacy. *Frontiers in Psychiatry* 14 (2023), 1080770.
- [26] Azadeh Forghani and Carman Neustaedter. 2014. The routines and needs of grandparents and parents for grandparent-grandchild conversations over distance. In *Proceedings of the sigchi conference on human factors in computing systems*. 4177–4186.
- [27] Yue Fu, Sami Foell, Xuhai Xu, and Alexis Hiniker. 2023. From Text to Self: Users’ Perceptions of Potential of AI on Interpersonal Communication and Self. *arXiv preprint arXiv:2310.03976* (2023).
- [28] Yue Fu, Sami Foell, Xuhai Xu, and Alexis Hiniker. 2024. From Text to Self: Users’ Perception of AIMC Tools on Interpersonal Communication and Self. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–17.
- [29] Verena Fuchsberger, Janne Mascha Beuthel, Philippe Bentégeac, and Manfred Tscheligi. 2021. Grandparents and grandchildren meeting online: the role of material things in remote settings. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–14.
- [30] Yumei Gan, Christian Greiffenhagen, and Stuart Reeves. 2020. Connecting distributed families: Camera work for three-party mobile video calls. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–12.
- [31] Narayan H Gandedkar, Matthew T Wong, and M Ali Darendeliler. 2021. Role of virtual reality (VR), augmented reality (AR) and artificial intelligence (AI) in tertiary education and research of orthodontics: An insight. In *Seminars in Orthodontics*, Vol. 27. Elsevier, 69–77.
- [32] Howard Giles, Kimberly A Noels, Angie Williams, Hiroshi Ota, Tae-Seop Lim, Sik Hung Ng, Ellen B Ryan, and Lilnabeth Somera. 2003. Intergenerational communication across cultures: Young people’s perceptions of conversations with family elders, non-family elders and same-age peers. *Journal of cross-cultural gerontology* 18 (2003), 1–32.
- [33] Jennifer Herron. 2016. Augmented reality in medical education and training. *Journal of Electronic Resources in Medical Libraries* 13, 2 (2016), 51–55.
- [34] Yasamin Heshmat and Carman Neustaedter. 2021. Family and friend communication over distance in Canada during the COVID-19 pandemic. In *Designing Interactive Systems Conference 2021*. 1–14.
- [35] Yasamin Heshmat, Carman Neustaedter, Kyle McCaffrey, William Odom, Ron Wakkary, and Zikun Yang. 2020. Familystories: Asynchronous audio storytelling for family members across time zones. In *Proceedings of the 2020 chi conference on human factors in computing systems*. 1–14.
- [36] Emil Rosenlund Hoeg, Jon Ram Bruun-Pedersen, Shannon Cheary, Lars Koreska Andersen, Razvan Paisa, Stefania Serafin, and Belinda Lange. 2021. Buddy biking: a user study on social collaboration in a virtual reality exergame for rehabilitation. *Virtual Reality* (2021), 1–18.
- [37] Kevin Huang, Patrick J. Sparto, Sara Kiesler, Asim Smailagic, Jennifer Mankoff, and Dan Siewiorek. 2014. A Technology Probe of Wearable In-Home Computer-Assisted Physical Therapy. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) (CHI ’14). Association for Computing Machinery, New York, NY, USA, 2541–2550. <https://doi.org/10.1145/2556288.2557416>
- [38] Hilary Hutchinson, Wendy Mackay, Bo Westerlund, Benjamin B Bederson, Allison Druin, Catherine Plaisant, Michel Beaudouin-Lafon, Stéphane Conversy, Helen Evans, Heiko Hansen, et al. 2003. Technology probes: inspiring design for and with families. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 17–24.

- [39] Katherine Isbister, Hideyuki Nakanishi, Toru Ishida, and Cliff Nass. 2000. Helper Agent: Designing an Assistant for Human-Human Interaction in a Virtual Meeting Space. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (The Hague, The Netherlands) (CHI '00). Association for Computing Machinery, New York, NY, USA, 57–64. <https://doi.org/10.1145/332040.332407>
- [40] Zhiqiu Jiang, Mashrur Rashik, Kunjal Panchal, Mahmood Jasim, Ali Sarvghad, Pari Riahi, Erica DeWitt, Fey Thurber, and Narges Mahyar. 2023. CommunityBots: Creating and Evaluating A Multi-Agent Chatbot Platform for Public Input Elicitation. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–32.
- [41] Qiao Jin, Ye Yuan, and Svetlana Yarosh. 2023. Socio-technical Opportunities in Long-Distance Communication Between Siblings with a Large Age Difference. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–15.
- [42] Xiaofu Jin, Wai Tong, Xiaoying Wei, Xian Wang, Emily Kuang, Xiaoyu Mo, Huamin Qu, and Mingming Fan. 2024. Exploring the Opportunity of Augmented Reality (AR) in Supporting Older Adults to Explore and Learn Smartphone Applications. 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 21, 18 pages. <https://doi.org/10.1145/3613904.3641901>
- [43] Bumsoo 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*. 1–16.
- [44] Bumsoo Kang, Seungwoo Kang, and Inseok Hwang. 2023. AI-driven Family Interaction Over Melded Space and Time. *IEEE Pervasive Computing* 22, 1 (2023), 85–94.
- [45] Candace L Kemp. 2005. Dimensions of grandparent-adult grandchild relationships: From family ties to intergenerational friendships. *Canadian Journal on Aging/La Revue canadienne du vieillissement* 24, 2 (2005), 161–177.
- [46] Susan Kemper. 2000. Accommodations to Aging. *Communication, technology and aging: Opportunities and challenges for the future* (2000), 30.
- [47] Jeong Hyun Kim, Jungkeun Kim, Jooyoung Park, Changju Kim, Jihoon Jhang, and Brian King. 2023. When ChatGPT gives incorrect answers: the impact of inaccurate information by generative AI on tourism decision-making. *Journal of Travel Research* (2023), 00472875231212996.
- [48] D Lawrence Kincaid. 2013. *Communication theory: Eastern and Western perspectives*. Academic Press.
- [49] Takeshi Kojima, Shixiang Shane Gu, Machel Reid, Yutaka Matsuo, and Yusuke Iwasawa. 2023. Large Language Models are Zero-Shot Reasoners. [arXiv:2205.11916](https://arxiv.org/abs/2205.11916) [cs.CL]
- [50] Anya Kolesnichenko, Joshua McVeigh-Schultz, and Katherine Isbister. 2019. Understanding emerging design practices for avatar systems in the commercial social vr ecology. In *Proceedings of the 2019 on Designing Interactive Systems Conference*. 241–252.
- [51] Marc Erich Latoschik, Daniel Roth, Dominik Gall, Jascha Achenbach, Thomas Waltemate, and Mario Botsch. 2017. The effect of avatar realism in immersive social virtual realities. In *Proceedings of the 23rd ACM symposium on virtual reality software and technology*. 1–10.
- [52] Sueyoon Lee, Abdallah El Ali, Maarten Wijntjes, and Pablo Cesar. 2022. Understanding and Designing Avatar Biosignal Visualizations for Social Virtual Reality Entertainment. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–15.
- [53] Yoonjoo Lee, Tae Soo Kim, Minsuk Chang, and Juho Kim. 2022. Interactive Children’s Story Rewriting Through Parent-Children Interaction. In *Proceedings of the First Workshop on Intelligent and Interactive Writing Assistants (In2Writing 2022)*. 62–71.
- [54] Jie Li, Yiping Kong, Thomas Röggl, Francesca De Simone, Swamy Ananthanarayan, Huib De Ridder, Abdallah El Ali, and Pablo Cesar. 2019. Measuring and understanding photo sharing experiences in social virtual reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–14.
- [55] Xiaoyan Li, Naomi Yamashita, Wen Duan, Yoshinari Shirai, and Susan R Fussell. 2023. Improving Non-Native Speakers’ Participation with an Automatic Agent in Multilingual Groups. *Proceedings of the ACM on Human-Computer Interaction* 7, GROUP (2023), 1–28.
- [56] 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 (2023), 1–26.
- [57] Weixin Liang, Girmaw Abebe Tadesse, Daniel Ho, Li Fei-Fei, Matei Zaharia, Ce Zhang, and James Zou. 2022. Advances, challenges and opportunities in creating data for trustworthy AI. *Nature Machine Intelligence* 4, 8 (2022), 669–677.
- [58] Mei-Chen Lin, Jake Harwood, and Jaye L Bonnesen. 2002. Conversation topics and communication satisfaction in grandparent-grandchild relationships. *Journal of Language and Social Psychology* 21, 3 (2002), 302–323.
- [59] Siân E Lindley, Richard Harper, and Abigail Sellen. 2009. Desiring to be in touch in a changing communications landscape: attitudes of older adults. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.

- 1693–1702.
- [60] Haochen Liu, Yiqi Wang, Wenqi Fan, Xiaorui Liu, Yaxin Li, Shaili Jain, Yunhao Liu, Anil Jain, and Jiliang Tang. 2022. Trustworthy ai: A computational perspective. *ACM Transactions on Intelligent Systems and Technology* 14, 1 (2022), 1–59.
 - [61] Michael Luck and Ruth Aylett. 2000. Applying artificial intelligence to virtual reality: Intelligent virtual environments. *Applied artificial intelligence* 14, 1 (2000), 3–32.
 - [62] Francisco Lopez Luro and Veronica Sundstedt. 2019. A comparative study of eye tracking and hand controller for aiming tasks in virtual reality. In *Proceedings of the 11th ACM Symposium on eye tracking research & applications*. 1–9.
 - [63] Kathleen M MacQueen, Eleanor McLellan, Kelly Kay, and Bobby Milstein. 1998. Codebook development for team-based qualitative analysis. *Cam Journal* 10, 2 (1998), 31–36.
 - [64] Torsten Maier, Nicolas F Soria Zurita, Elizabeth Starkey, Daniel Spillane, Christopher McComb, and Jessica Menold. 2022. Comparing human and cognitive assistant facilitated brainstorming sessions. *Journal of Engineering Design* 33, 4 (2022), 259–283.
 - [65] Karl Manheim and Lyric Kaplan. 2019. Artificial intelligence: Risks to privacy and democracy. *Yale JL & Tech.* 21 (2019), 106.
 - [66] Kleanthis Manolakis and George Papagiannakis. 2022. Virtual Reality simulation streamlines medical training for healthcare professionals. *Journal of dentistry* 121 (2022), 103987.
 - [67] Nora McDonald, Sarita Schoenebeck, and Andrea Forte. 2019. Reliability and Inter-rater Reliability in Qualitative Research: Norms and Guidelines for CSCW and HCI Practice. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 72 (nov 2019), 23 pages. <https://doi.org/10.1145/3359174>
 - [68] Joshua McVeigh-Schultz, Anya Kolesnichenko, and Katherine Isbister. 2019. Shaping pro-social interaction in VR: an emerging design framework. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.
 - [69] Ines Miguel-Alonso, David Checa, Henar Guillen-Sanz, and Andres Bustillo. 2024. Evaluation of the novelty effect in immersive Virtual Reality learning experiences. *Virtual Reality* 28, 1 (2024), 27.
 - [70] Mirja Mittermaier, Mariam M Raza, and Joseph C Kvedar. 2023. Bias in AI-based models for medical applications: challenges and mitigation strategies. *NPJ Digital Medicine* 6, 1 (2023), 113.
 - [71] Fabiane FR Morgado, Juliana FF Meireles, Clara M Neves, Ana Amaral, and Maria EC Ferreira. 2017. Scale development: ten main limitations and recommendations to improve future research practices. *Psicologia: Reflexão e Crítica* 30 (2017).
 - [72] Fares Moustafa and Anthony Steed. 2018. A longitudinal study of small group interaction in social virtual reality. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*. 1–10.
 - [73] Elizabeth D Mynatt, Jim Rowan, Sarah Craighill, and Annie Jacobs. 2001. Digital family portraits: supporting peace of mind for extended family members. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 333–340.
 - [74] Ayesha Nadeem, Babak Abedin, and Olivera Marjanovic. 2020. Gender bias in AI: A review of contributing factors and mitigating strategies. (2020).
 - [75] Carman Neustaedter, Kathryn Elliot, and Saul Greenberg. 2006. Interpersonal awareness in the domestic realm. In *Proceedings of the 18th Australia conference on Computer-Human Interaction: Design: Activities, Artefacts and Environments*. 15–22.
 - [76] Alex Nichol, Heewoo Jun, Prafulla Dhariwal, Pamela Mishkin, and Mark Chen. 2022. Point-e: A system for generating 3d point clouds from complex prompts. *arXiv preprint arXiv:2212.08751* (2022).
 - [77] Nicholas R Nicholson. 2012. A review of social isolation: an important but underassessed condition in older adults. *The journal of primary prevention* 33, 2 (2012), 137–152.
 - [78] Eirini Ntoutsis, Pavlos Fafalios, Ujwal Gadiraju, Vasileios Iosifidis, Wolfgang Nejdl, Maria-Esther Vidal, Salvatore Ruggieri, Franco Turini, Symeon Papadopoulos, Emmanouil Krasanakis, et al. 2020. Bias in data-driven artificial intelligence systems—An introductory survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 10, 3 (2020), e1356.
 - [79] Soo Youn Oh, Jeremy Bailenson, Erika Weisz, and Jamil Zaki. 2016. Virtually old: Embodied perspective taking and the reduction of ageism under threat. *Computers in Human Behavior* 60 (2016), 398–410.
 - [80] Bhikizizwe Peterson. 2019. Spectrality and inter-generational black narratives in South Africa. *Social Dynamics* 45, 3 (2019), 345–364.
 - [81] Alisha Pradhan, Ben Jelen, Katie A Siek, Joel Chan, and Amanda Lazar. 2020. Understanding older adults’ participation in design workshops. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–15.
 - [82] Hayes Raffle, Glenda Revelle, Koichi Mori, Rafael Ballagas, Kyle Buza, Hiroshi Horii, Joseph Kaye, Kristin Cook, Natalie Freed, Janet Go, et al. 2011. Hello, is grandma there? let’s read! StoryVisit: family video chat and connected e-books. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 1195–1204.

- [83] Luiz Rodrigues, Filipe D Pereira, Armando M Toda, Paula T Palomino, Marcela Pessoa, Leandro Silva Galvão Carvalho, David Fernandes, Elaine HT Oliveira, Alexandra I Cristea, and Seiji Isotani. 2022. Gamification suffers from the novelty effect but benefits from the familiarization effect: Findings from a longitudinal study. *International Journal of Educational Technology in Higher Education* 19, 1 (2022), 13.
- [84] Natalia Romero, Panos Markopoulos, Joy Van Baren, Boris De Ruyter, Wijnand Ijsselstein, and Babak Farshchian. 2007. Connecting the family with awareness systems. *Personal and Ubiquitous Computing* 11, 4 (2007), 299–312.
- [85] Drew Roselli, Jeanna Matthews, and Nisha Talagala. 2019. Managing bias in AI. In *Companion proceedings of the 2019 world wide web conference*. 539–544.
- [86] Daniel Roth, Constantin Klelnbeck, Tobias Feigl, Christopher Mutschler, and Marc Erich Latoschik. 2018. Beyond replication: Augmenting social behaviors in multi-user virtual realities. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 215–222.
- [87] Mikko Salminen, Simo Järvelä, Antti Ruonala, Janne Timonen, Kristiina Mannermaa, Niklas Ravaja, and Giulio Jacucci. 2018. Bio-adaptive social VR to evoke affective interdependence: DYNECOM. In *23rd international conference on intelligent user interfaces*. 73–77.
- [88] *Kelsea Schulenberg, Lingyuan Li, Guo Freeman, Samaneh Zamanifard, and Nathan J McNeese. 2023. Towards Leveraging AI-based Moderation to Address Emergent Harassment in Social Virtual Reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–17.
- [89] Alexander Seifert. 2020. The digital exclusion of older adults during the COVID-19 pandemic. *Journal of gerontological social work* 63, 6-7 (2020), 674–676.
- [90] Procheta Sen and Debasis Ganguly. 2020. Towards socially responsible ai: Cognitive bias-aware multi-objective learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34. 2685–2692.
- [91] Chenxinran Shen, Joanna McGrenere, and Dongwook Yoon. 2024. LegacySphere: Facilitating Intergenerational Communication Through Perspective-Taking and Storytelling in Embodied VR. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '24). Association for Computing Machinery, New York, NY, USA, Article 119, 16 pages. <https://doi.org/10.1145/3613904.3641923>
- [92] Harrison Jesse Smith and Michael Neff. 2018. Communication behavior in embodied virtual reality. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–12.
- [93] Laura Stafford. 2004. *Maintaining long-distance and cross-residential relationships*. Routledge.
- [94] StudyMassCom. 2023. Eastern vs. Western Perspective of Communication. <https://studymasscom.com/communication/eastern-and-western-perception-of-communication/> [Online; accessed 24-December-2023].
- [95] Luma Tabbaa, Ryan Searle, Saber Mirzaee Bafti, Md Moinul Hossain, Jitrapol Intarasirisawat, Maxine Glancy, and Chee Siang Ang. 2021. Vreed: Virtual reality emotion recognition dataset using eye tracking & physiological measures. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 5, 4 (2021), 1–20.
- [96] Alan C Taylor, Mihaela Robila, and Hae Seung Lee. 2005. Distance, contact, and intergenerational relationships: Grandparents and adult grandchildren from an international perspective. *Journal of adult development* 12, 1 (2005), 33–41.
- [97] Kimberly Tee, AJ Bernheim Brush, and Kori M Inkpen. 2009. Exploring communication and sharing between extended families. *International Journal of Human-Computer Studies* 67, 2 (2009), 128–138.
- [98] A. M. Tomlin. 1998. Grandparents' influence on grandchildren. In *Handbook on grandparenthood*. Westport, CT: Greenwood Press., 159–170.
- [99] Margarita Vinnikov, Robert S Allison, and Suzette Fernandes. 2017. Gaze-contingent auditory displays for improved spatial attention in virtual reality. *ACM Transactions on Computer-Human Interaction (TOCHI)* 24, 3 (2017), 1–38.
- [100] René Vutborg, Jesper Kjeldskov, Sonja Pedell, and Frank Vetere. 2010. Family storytelling for grandparents and grandchildren living apart. In *Proceedings of the 6th Nordic conference on human-computer interaction: Extending boundaries*. 531–540.
- [101] Chao-Ming Wang, Cheng-Hao Shao, and Cheng-En Han. 2022. Construction of a Tangible VR-Based Interactive System for Intergenerational Learning. *Sustainability* 14, 10 (2022), 6067.
- [102] Xiaoying Wei, Yizheng Gu, Emily Kuang, Xian Wang, Beiyang Cao, Xiaofu Jin, and Mingming Fan. 2023. Bridging the Generational Gap : Exploring How Virtual Reality Supports Remote Communication Between Grandparents and Grandchildren. In *CHI Conference on Human Factors in Computing Systems*.
- [103] Xiaoying Wei, Xiaofu Jin, and Mingming Fan. 2022. Communication in Immersive Social Virtual Reality: A Systematic Review of 10 Years' Studies. *arXiv preprint arXiv:2210.01365* (2022).
- [104] Angie Williams and Howard Giles. 1996. Intergenerational conversations: Young adults' retrospective accounts. *Human Communication Research* 23, 2 (1996), 220–250.
- [105] Angie Williams and Jon F Nussbaum. 2013. *Intergenerational communication across the life span*. Routledge.
- [106] Angie Williams, Hiroshi Ota, Howard Giles, Herbert D Pierson, Cynthia Gallois, Sik-Hung Ng, Tae-Seop Lim, Ellen Bouchard Ryan, Lilnabeth Somera, John Maher, et al. 1997. Young people's beliefs about intergenerational

communication: An initial cross-cultural comparison. *Communication Research* 24, 4 (1997), 370–393.

[107] Wei Xu, Marvin J Dainoff, Liezhong Ge, and Zaifeng Gao. 2023. Transitioning to human interaction with AI systems: New challenges and opportunities for HCI professionals to enable human-centered AI. *International Journal of Human–Computer Interaction* 39, 3 (2023), 494–518.

[108] Jincai Yang, Cheng Shen, and Niu Huang. 2020. Predicting or pretending: artificial intelligence for protein-ligand interactions lack of sufficiently large and unbiased datasets. *Frontiers in pharmacology* 11 (2020), 69.

[109] Svetlana Yarosh, Anthony Tang, Sanika Mokashi, and Gregory D Abowd. 2013. "almost touching" parent-child remote communication using the sharetable system. In *Proceedings of the 2013 conference on Computer supported cooperative work*. 181–192.

[110] Yan Bing Zhang and Mary Lee Hummert. 2001. Harmonies and tensions in Chinese intergenerational communication: Younger and older adults' accounts. *Journal of Asian Pacific Communication* 11, 2 (2001), 203–230.

[111] Zheng Zhang, Ying Xu, Yanhao Wang, Bingsheng Yao, Daniel Ritchie, Tongshuang Wu, Mo Yu, Dakuo Wang, and Toby Jia-Jun Li. 2022. Storybuddy: A human-ai collaborative chatbot for parent-child interactive storytelling with flexible parental involvement. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–21.

[112] 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 *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–24.

A Appendix

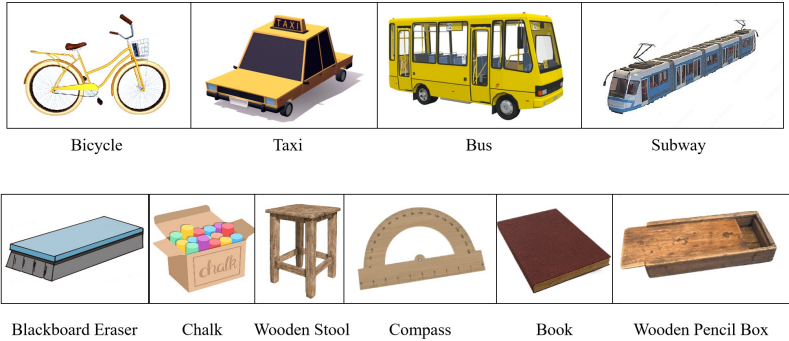


Fig. 5. Pre-generated 3D objects

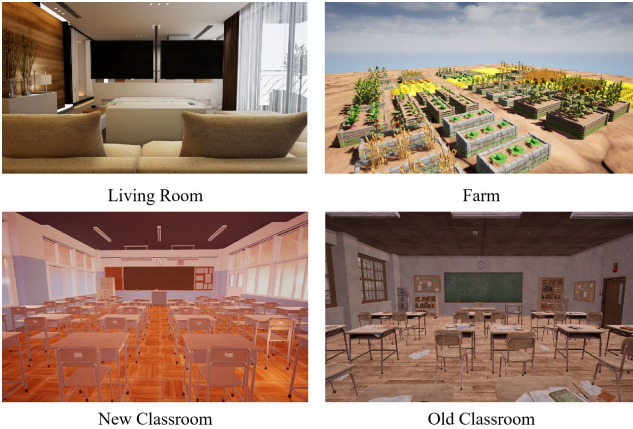


Fig. 6. Pre-generated virtual scenes

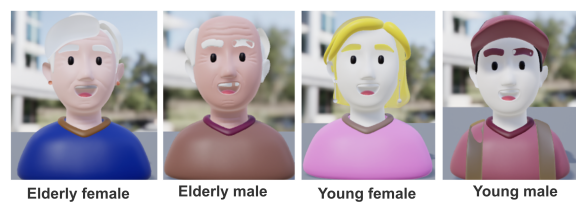


Fig. 7. Four types of avatars

Table 3. Summary of our probes, used prompts in each probe, purposes, and the number of groups using this probe for specific purposes.

Probe	Prompts	Purpose	Group Usage
Image Generation	Click a button and say: <i>“Please give me a picture of ...”</i>	• Generated images of attractions and food to fill out the timetable in Task 1.	12
		• Generated images of attractions unfamiliar to the other generation to negotiate travel plans in Task 1.	3
		• Generated images of campus facilities to better share their campus life in Task 2.	7
		• Generated images of working places to describe their current jobs and future aspirations.	2
3D Objects Generation	Click a button and say: <i>“Please give me ... (e.g., “Bicycle”, “Pencil box”, “Chalk”, etc...)”</i>	• Generated 3D models of vehicles to fill out the timetable in Task 1.	12
		• Generated 3D models of stationery items to share campus life in Task 2.	7
Scene Generation	Click a button and say: <i>“Please take me to ... (e.g., “farm”, “old school”, “old classroom”, etc...)”</i>	• Engaged in farm or old classrooms scene in VR to reminiscence and assist storytelling.	10
		• Switched from new and old classrooms in VR to discuss and compare their school life.	6
Info Assistant	Click a button and verbally describe the needs	• Asked for AI suggestion of travel plan in Task 1.	9
		• Queried information that both parties are not familiar with (e.g., “Where is this section of the Great Wall located?”).	2
		• Sought authoritative and systematic explanations of specific words (e.g., “nèi juǎn”).	2
Chat Facilitator	[Automatically] provide intervention based on communication contexts to provide appropriate voice suggestions	• Suggested topics based on contextual content, helping two generations avoid awkward silences	8
		• Provided timely interventions to maintain a balanced conversation	6
Emotional Visualization	[Automatically] activate when detecting keywords	• Helped inter-generational participants in expressing their emotions more effectively.	7

Received January 2024; revised July 2024; accepted October 2024