

SimViews: An Interactive Multi-Agent System Simulating Visitor-to-Visitor Conversational Patterns to Present Diverse Perspectives of Artifacts in Virtual Museums

Mingyang Su*

The Hong Kong University of Science
and Technology (Guangzhou)
Guangzhou, China
sumy22@mails.tsinghua.edu.cn

Chao Liu*

The Hong Kong University of Science
and Technology (Guangzhou)
Guangzhou, China
cliu009@connect.hkust-gz.edu.cn

Jingling Zhang

The Hong Kong University of Science
and Technology (Guangzhou)
Guangzhou, China
jzhang898@connect.hkust-gz.edu.cn

Wu Shuang

The Hong Kong University of Science
and Technology (Guangzhou)
Guangzhou, China
shuang@u.nus.edu

Mingming Fan†

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

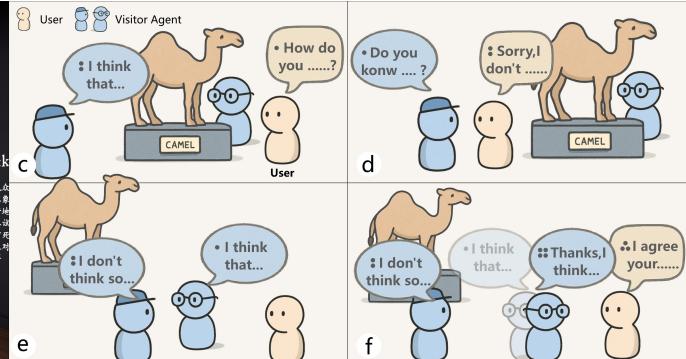


Figure 1: (a) Multi-agent conversation in the virtual museum to present diverse perspectives; (b) The participant engages in a conversation with the visitor agent; Four types of conversation patterns: (c) The user's active speaking: the user actively initiates a conversation with the agent; (d) The user's passive speaking: the agent initiates a conversation with the user, and the user responds; (e) The user's passive listening: the user overhears the agents' conversation; (f) The user's active listening: the user listens to the agents' conversation, and they join in if they want.

Abstract

Offering diverse perspectives on a museum artifact can deepen visitors' understanding and help avoid the cognitive limitations of a single narrative, ultimately enhancing their overall experience.

*Both authors contributed equally to this research.

†Corresponding author

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.

MM '25, Dublin, Ireland

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

ACM ISBN 979-8-4007-2035-2/2025/10

<https://doi.org/10.1145/3746027.3754593>

Physical museums promote diversity through visitor interactions. However, it remains a challenge to present multiple voices appropriately while attracting and sustaining a visitor's attention in the virtual museum. Inspired by recent studies that show the effectiveness of LLM-powered multi-agents in presenting different opinions about an event, we propose SimViews, an interactive multi-agent system that simulates visitor-to-visitor conversational patterns to promote the presentation of diverse perspectives. The system employs LLM-powered multi-agents that simulate virtual visitors with different professional identities, providing diverse interpretations of artifacts. Additionally, we constructed 4 conversational patterns between users and agents to simulate visitor interactions. We conducted a within-subject study with 20 participants, comparing SimViews to a traditional single-agent condition.

Our results show that SimViews effectively facilitates the presentation of diverse perspectives through conversations, enhancing participants' understanding of viewpoints and engagement within the virtual museum.

CCS Concepts

- Human-centered computing → *Interaction design; Collaborative interaction.*

Keywords

Multi-agent, diverse perspective, virtual museum

ACM Reference Format:

Mingyang Su, Chao Liu, Jingling Zhang, Wu Shuang, and Mingming Fan. 2025. SimViews: An Interactive Multi-Agent System Simulating Visitor-to-Visitor Conversational Patterns to Present Diverse Perspectives of Artifacts in Virtual Museums. In *Proceedings of the 33rd ACM International Conference on Multimedia (MM '25), October 27–31, 2025, Dublin, Ireland*. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3746027.3754593>

1 Introduction

A long-standing metaphor for the museum is a ‘forum’ for different voices to debate and converse [9, 37]. Presenting diverse perspectives in museums can help visitors move beyond personal biases, fostering more inclusive and meaningful experiences [10, 15, 67]. Specifically, it can deepen understanding [11, 12, 61], encourage cultural reflection [1], and enhance engagement [22, 53]. In physical museums, such diversity is typically promoted through participatory practices, such as inviting underrepresented groups to contribute to exhibition narratives during curation [40], or collecting visitor feedback through comment books and cards [14, 52].

With the rapid advancement of digital technology, virtual museums are currently contributing to cultural diversity. However, these diversities focus more on the content of exhibits, purposes of usage, and access [25], ignoring the presentation of diverse perspectives within the museum. Researchers suggest the use of AI practice to achieve this goal and provide a more inclusive service [31].

Recent studies have shown that Large Language Models (LLMs) can effectively simulate agents within a given domain [41, 59]. These agents are widely used in social media [80], online decision-making [57], strategic consulting [54], and other areas to provide diverse voices through dialogue, which may offer the potential to display diverse perspectives in virtual museums. However, most of these are primarily text-based, the museum should actively encourage and facilitate different forms of conversation [63], rather than being confined to a static and monologic wall text.

Inspired by these works, we explored *how to design an LLM-powered multi-agent conversation method to effectively present diverse perspectives through conversation with users? (RQ1)*. To this end, we developed SimViews, an interactive system that simulates visitor-to-visitor conversational patterns to support the presentation of diverse perspectives in virtual museums. The system uses LLMs to create visitor agents with distinct professional identities based on literature [80], and structures 4 visitor-to-visitor conversation patterns between users and agents by drawing on Goffman’s participation framework [24]: **active speaking** (user initiates), **passive speaking** (agent initiates), **passive listening** (user overhears

agent conversations), and **active listening** (user joins the ongoing discussion of two agents), as shown in Fig. 1.

We also explored *whether and how multi-agent conversations supported by SimViews may help users understand these diverse perspectives and engage more deeply with the museum? (RQ2)*. To address this, we implemented a prototype of a virtual museum featuring two exhibits and six agents. We then conducted a user study with 20 participants. The results showed that SIMVIEWS effectively facilitates the presentation of diverse perspectives through conversations, enhancing participants' understanding of viewpoints and engagement within the virtual museum. Among the four patterns, participants rated **active listening** as most helpful for understanding and engagement, while **active speaking** was preferred for gaining control. Based on these findings, we present design insights for presenting diverse perspectives in a virtual museum through multi-agent conversation design. Overall, our work made the following contributions:

- We propose a multi-agent system that simulates 4 visitor-to-visitor conversational patterns to present diverse perspectives in virtual museums.
- We implemented a functional prototype and conducted a user study to evaluate the design, offering design insights into presenting multiple perspectives in virtual museums using multi-agent systems.

2 Related Work

2.1 Diverse Museum

To embrace and foster cultural diversity, ICOM redefined museums as “democratizing, inclusive, and polyphonic spaces” [33], encouraging the inclusion of diverse voices from outside the institution in exhibit interpretation [61]. In physical museums, such diversity is increasingly realized through curatorial practices that involve underrepresented groups, particularly minority communities, which sometimes includes seeking interpretations from the communities who created, used, or are associated with them [40, 68]. Visitor participation also plays a key role in presenting multiple perspectives during museum visits [61]. Many museums now employ interactive installations [52] and multimedia technologies [58] to support this goal. For example, Tate Britain’s Turner Prize exhibition includes a space for visitors to share their responses, though these are often limited to text-based displays. At Randers Art Museum, visitors were invited to comment on artworks, and their reflections were later exhibited, fostering perspective diversity through interaction and dialogue [63].

Virtual museums also contribute to cultural diversity, primarily by broadening access to diverse audiences and offering a wide range of content [25, 74]. However, this form of diversity often overlooks the presentation of multiple perspectives on the artifacts themselves, which remains a significant gap in the virtual museum experience.

2.2 LLM-powered Multi-agent

LLMs have demonstrated remarkable capabilities in simulating domain-specific agents through coherent prompting and in-context learning [41, 59]. By maintaining consistent personas and delivering context-aware responses, LLMs can effectively embody expert

roles over extended interactions [47, 76]. Building on this capability, recent studies have introduced LLM-based agents into virtual museum environments, where they serve as personalized tour guides. These agents can offer professional, adaptive explanations and contextual guidance tailored to users' interests [71, 72, 74].

To further enrich interaction, researchers have explored multi-agent frameworks that leverage LLMs' reasoning through inter-agent collaboration. Multi-agent setups allow agents to discuss, debate, and converge on more accurate or nuanced answers [19, 27, 78, 81]. For instance, multiple expert agents can jointly analyze medical cases to reach a diagnostic consensus [69]. Beyond accuracy, multi-agent systems also introduce identity diversity, agents can be designed to represent different expertise, personalities, or viewpoints, thereby reducing the uniformity of responses typically generated by a single LLM [8, 43]. One example is Zhang et al.'s system that simulates diverse agents with distinct professional backgrounds during social media browsing to expose users to a broader range of perspectives [80].

2.3 Agent Conversation

Benefiting from LLMs' powerful understanding and generation capabilities, multi-agents can engage in more natural interactions with humans. However, in current studies, user interactions with multi-agents are still primarily based on dialogue boxes [21, 69, 80]. For example, the Finch financial advisory system enables users to make decisions through text-based conversations with four collaborative agents [16, 17]. While it showcases the potential of multi-agent collaboration in complex tasks, it offers limited immersion and interactivity in virtual environments.

To improve user experience, recent studies have explored virtual avatars and voice interaction to support more natural conversations [46, 49, 70]. For instance, HUMAINE introduced voice interaction in simulated transaction negotiations [18], but its agents only engage in dialogue with users, lacking inter-agent discussions. Agent United provides a multi-agent dialogue platform, where although agents are equipped with virtual avatars, users can only participate through text-based response options [3], limiting both flexibility and realism. These examples highlight that achieving truly natural and flexible multi-agent conversation remains a significant research challenge.

3 System Design

To address RQ1, we proposed key design considerations and developed a multi-agent system that incorporates visitor agent identities and structured conversation patterns to support inclusive and engaging visitor interactions.

3.1 Agent Design

For agent design, we addressed two primary considerations: **(D1.1) Design virtual agents with distinct professional identities.** Museum visitors often have multiple social identities, leading to diverse interpretations of exhibits [65]. For example, historians viewing a Bengal tiger specimen from India in an English museum might interpret it as "a cipher for hegemonic Victorian attitudes toward India and its people" [13], whereas animal conservationists could emphasize ethical objections to hunting practices [26].

To reflect this interpretive diversity, our agents were created with specific professional identities instead of generic visitor personas. **(D1.2) Equip agents with multimodal representations, including visual avatars and anthropomorphic voices.** Museums need interactive presentations beyond static, text-based displays to effectively engage visitors [63]. Prior studies highlight that multimodal forms of information significantly enhance participant engagement and comprehension in virtual environments [44, 60]. Additionally, agents with visual avatars and natural voices are more persuasive and engaging [2]. Therefore, our agent representations include detailed avatar models and voice interactions instead of mere textual dialogue boxes.

3.2 Conversation Design

For conversation design, our goal is to simulate the visitor-to-visitor conversational patterns. Inspired by recent studies that have applied Goffman's participation framework [20, 24] to multi-agent dialogue settings [64, 73], we adopt this framework to guide our design and identified two key dimensions for structuring visitor-to-visitor conversational patterns: **(D2.1) Define user roles as speaker or listener.** In multi-agent environments, users may act as speakers, addressed listeners, or bystanders. Differentiating these roles allows for the design of varied and realistic conversational interactions. **(D2.2) Specify whether the user or the agent initiates the conversation.** Clarifying conversational initiative helps determine interaction flow and user engagement levels, guiding how agents are scripted to respond or lead.

3.3 Design Features

3.3.1 Professionally Grounded Agent Identities. To address the agent identity requirement **(D1.1)**, we assigned each visitor agent a distinct professional identity based on the scholarly perspectives of exhibits, such as ethicists, art historians, and biologists. Prior studies have shown that simulating conversations across professional roles enables users to access more diverse perspectives [80]. However, unlike open-ended online platforms, museums demand more professional and accurate representations [38, 56]. To reduce hallucinations and ensure credibility, each agent's identity and dialogue were grounded in scholarly literature, using content authored by domain experts. This design enables the system to present reliably diverse and academically sound interpretations of the exhibits.

3.3.2 Embodied and Multimodal Agent Representation. To fulfill representation considerations **(D1.2)**, we provided each agent with a full-body 3D avatar and a synthesized voice to facilitate natural voice-based interactions. This multimodal embodiment goes beyond traditional dialogue boxes and enhances immersion. Research has shown that voice and visual embodiment increase believability, trust, and engagement in virtual agents [2, 60]. Identity labels are hidden by default and shown only when users interact with an agent, helping participants distinguish agent roles while minimizing cognitive load and visual clutter. To simulate lifelike museum dynamics, agents autonomously navigate the gallery, observe exhibits, and interact based on behavior tree logic. As shown in Fig. 2, when a user enters, agents exhibit socially present behaviors such as walking, gesturing, and conversing.

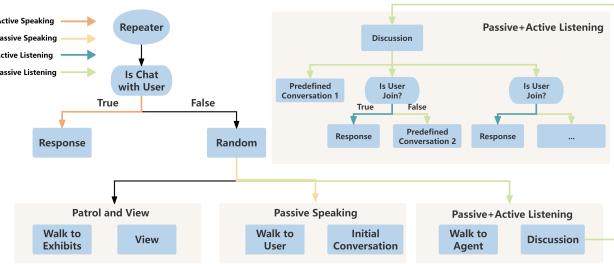


Figure 2: Agent behavior tree for switching between patrol, viewing, and four conversational patterns.

3.3.3 Multi-Pattern Conversational Framework. Based on the identified conversation dimensions (D2.1 and D2.2), we proposed a 2×2 matrix to categorize different patterns of user-agent conversations. This structure yields four conversation types: **Active speaking**: The user initiates the conversation, and the agent responds. This is the most common interaction pattern in traditional user-agent systems and virtual environments, such as when a visitor asks a guide agent a question [74]. **Passive speaking**: The agent initiates the conversation, and the user responds. This reflects recent attention to agent proactivity and system-driven engagement strategies [48]. **Active listening**: The user listens to an ongoing conversation between agents and may choose to join. This corresponds to an eavesdropping-like behavior where the user transitions from passive observer to active participant. **Passive listening**: the user overhears an agent conversation without actively engaging. This mode represents ambient awareness or unintended reception of dialogue, aligning with Goffman's notion of overhearers.

4 SimViews Prototype Implementation

Building on our design features, we developed **SIMVIEWS** as a working prototype that brings multi-agent visitor conversations into a virtual environment.

4.1 Virtual Museum Setup

We built a virtual gallery in Unity and selected two exhibits with contrasting interpretations from prior literature: the sculpture *Lion Attacking a Dromedary* and the performance artwork *The Artifact Piece*. The former depicts a dramatic animal attack [55], while the latter is a self-exhibition by Native American artist James Luna challenging ethnographic display norms [28]. These exhibits were selected to maximize interpretive contrast, thematic richness, and scholarly coverage. Both have been widely discussed from multiple disciplinary perspectives, making them suitable anchors for multi-agent conversations. To ensure visual fidelity, we invited an animator to reconstruct both exhibits in 3D based on photographs. Lighting, textures, and materials were carefully refined, and exhibit descriptions were displayed within the environment (see Fig. 3).

4.2 Constructing Multiple Visitor Agents

We extracted three distinct viewpoints for each exhibit from relevant literature and assigned them to separate agents to ensure a clear and coherent expression of diverse perspectives. For *Lion*

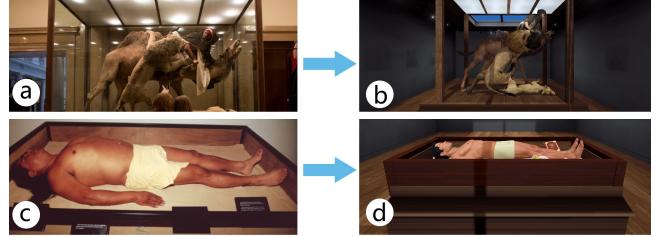


Figure 3: Original and reconstructed versions of *Lion Attacking a Dromedary* (a–b) and *The Artifact Piece* (c–d).

Attacking a Dromedary, the viewpoints include: (1) an emphasis on the sculpture's Romantic narrative and visual drama; (2) a critique of the use of human remains, raising concerns about racial violence; and (3) a challenge to the depiction of nature as inherently violent, advocating for a more balanced ecological perspective [4, 23, 55]. These viewpoints were assigned to three different agents. Then, we assigned corresponding identity labels to these agents based on these viewpoints, in this case **Aesthete**, **Ethicist**, and **Biologist**. For *The Artifact Piece*, we adopted the same approach and assigned the extracted perspectives to **Art Historian**, **Indigenous Scholar**, and **Curator**, based on their literature [5, 6, 28, 39].

Each agent features a unique appearance, identity, and conversational capability. For agent representation, avatars were generated using Ready Player Me, with randomized attributes (e.g., gender, skin tone, and clothing). To reduce stereotypes, agent appearance was not linked to specific professions. Identity labels were hidden until interaction began, ensuring that participants focused on the agents' conversational content rather than visual cues. Character animations such as standing, walking, talking, and thinking were integrated using Unity's Animator. To introduce variability, Unity's Behavior Tree and Animator were used to randomize these actions. Each agent's voice was synthesized using Azure's text-to-speech service, matching the avatar gender. Lip-sync was implemented via BlendShapes and synchronized with speech output.

4.3 Multi-agent conversational flow

The system integrates agent-agent and user-agent conversation flows within a real-time virtual environment (see Fig. 4). To enable autonomous multi-agent interaction, each agent is programmed to initiate or respond to spoken interaction based on role and proximity. Agent-to-agent interactions are driven by predefined scripts drawn from literature to present contrasting perspectives. These conversations are coordinated through behavior trees, as shown in Fig. 2, ensuring narrative coherence and accuracy. Agents also autonomously perform actions such as walking, observing exhibits, initiating conversations, or engaging in discussion. When interacting with users, the system transcribes speech using Azure's speech-to-text service, combines it with exhibit context and literature-based viewpoints, and sends it as a prompt to the Spark LLM. The generated response is then synthesized into speech using Azure TTS, with lip movement and hand gestures automatically synchronized. These mechanisms create a lifelike and socially engaging environment that simulates visitor-to-visitor conversations in physical museums.

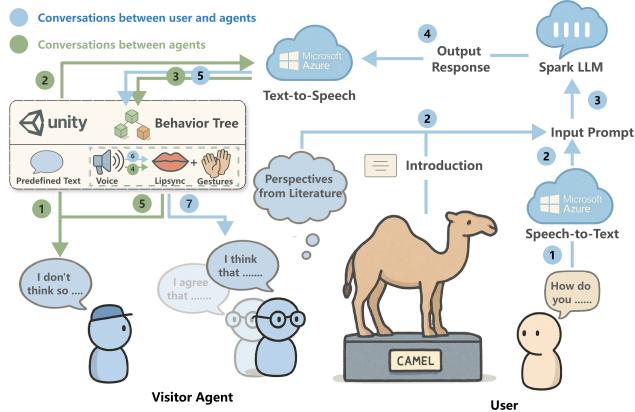


Figure 4: Interaction flow of a multi-agent system supporting voice-based user-agent and agent-agent conversations.

5 USER STUDY

5.1 Study Design

To address RQ2, we conducted a within-subject comparative experiment. The experiment included two conditions, illustrated in Fig. 5. In the **SIMVIEWS** condition, participants interacted with multiple agents, each representing a distinct expert identity. In the **BASE** condition, participants interacted with a single guide agent, which presented all viewpoints in a unified narrative, sequentially explaining perspectives based on the same literature sources as those used in **SIMVIEWS**. Participants could ask follow-up questions and receive direct responses. Each participant experienced both conditions with different exhibits, and the order was counterbalanced using a Latin Square design. This study was reviewed and approved by the ethics committee at the first author's institution.

5.2 Participants

We recruited 20 participants ($M=25.15$, $SD=2.43$, aged 18–34) through word-of-mouth and snowball sampling. To minimize bias in understanding diverse perspectives, we recorded their professional backgrounds and museum visit frequency. Participants represented diverse disciplines: 4 from Computer Science, 3 from Environmental Studies, 3 from Design, 2 each from Economics, Materials Science, Biomedical Sciences, and Social Sciences, and 1 each from Chemistry and Intelligent Manufacturing. In terms of visit habits, 40% visited semiannually, followed by 30% quarterly, and 15% each for monthly and annual visits.

5.3 Procedure

Before the study began, participants completed a practice session using the Leeds Tiger exhibit, presented in a separate gallery. This exhibit, not used in the main study, allowed participants to familiarize themselves with the interaction modes of both **SIMVIEWS** and **BASE**, with researcher guidance provided throughout. All interactions were recorded. Each participant spent at least 10 minutes in each exhibit, and their behavior was recorded. After each condition, participants completed a knowledge test, shared their viewpoints, and filled out a questionnaire. Researchers followed up with brief

questions and semi-structured interviews to gather deeper insights and design feedback.

To evaluate our systems, we collected four types of data: custom questionnaires, conversational data, knowledge tests, and viewpoint statements. Questionnaires captured participants' subjective experiences. Conversational data recorded user-agent interactions and were coded for dialogue frequency and type. Knowledge tests, consisting of 10 researcher-designed multiple-choice questions, and viewpoint statements were used to evaluate participants' understanding of diverse perspectives.

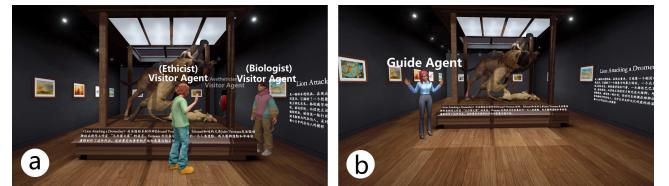


Figure 5: (a) In **SIMVIEWS: participants interact with visitor agents; (b) In **BASE**: participants listen to the guide agent or ask questions.**

6 Results

6.1 Custom Questionnaire

We first ran Shapiro-Wilk tests and found that the data were not normally distributed, so we used the Mann-Whitney U test for analysis. The results are shown in Fig. 6. To understand participants' preferences for different conversation patterns, we also asked them to rank four types across four dimensions. Rankings were converted to weighted scores (1st = 4, 2nd = 3, 3rd = 2, 4th = 1), and used the Kruskal-Wallis test. The results are shown in Fig. 7.

Understanding: Q1 (*I find it helpful for understanding the diverse perspectives on the exhibition more comprehensively.*) showed a significant difference ($U = 128.5$, $p < 0.05$), with the **SIMVIEWS** ($Md = 6.0$, $IQR = 2.0$) was higher than the **BASE** ($Md = 5.5$, $IQR = 2.0$). Similarly, Q2 (*I find it helpful for understanding the exhibition more deeply.*) also revealed a significant difference ($U = 129.5$, $p < 0.05$), with the **SIMVIEWS** ($Md = 6.0$, $IQR = 1.0$) outperforming the **BASE** ($Md = 5.0$, $IQR = 1.75$).

Engagement: Q3 (*I think it improved my engagement with the exhibition*) showed a significant difference ($U = 108.5$, $p < 0.05$), the **SIMVIEWS** ($Md = 6.0$, $IQR = 1.0$) was significantly higher than the **BASE** ($Md = 4.0$, $IQR = 1.75$). However, Q4 (*I am curious about the exhibits and actively explore their details*) did not reach significance ($U = 142.0$, $p = 0.099$).

Satisfaction: Q5 (*I feel it met my expectations.*) showed a significant difference ($U = 108.0$, $p < 0.01$), with the **SIMVIEWS** ($Md = 6.0$, $IQR = 1.75$) receiving higher ratings than the **BASE** ($Md = 5.0$, $IQR = 2.0$). Similarly, Q6 (*I hope to see it in more museums.*) also reached significance ($U = 110$, $p < 0.05$), with the **SIMVIEWS** ($Md = 6.5$, $IQR = 2.0$) again scoring higher than the **BASE** ($Md = 5.0$, $IQR = 2.0$).

Usefulness: Q7 (*I can easily get diverse perspectives through it.*) showed a significant difference ($U = 129$, $p < 0.05$), with participants rating the **SIMVIEWS** ($Md = 6.0$, $IQR = 1.0$) higher than

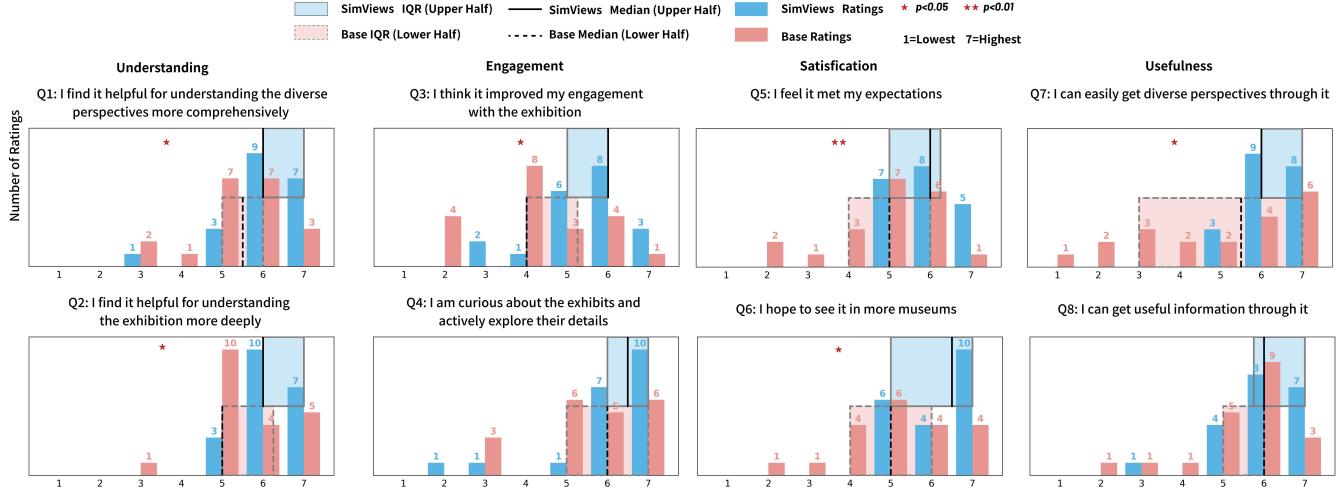


Figure 6: Results of our custom questionnaire on a Likert scale (1 = strongly disagree, 7 = strongly agree). Blue represents ratings for SIMVIEWS, while red represents ratings for BASE. The shaded areas indicate the interquartile range (IQR), and the vertical lines mark the median values for each group. Statistical significance is indicated as * $p < 0.05$, ** $p < 0.01$.

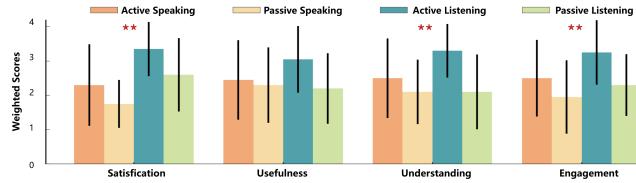


Figure 7: Weighted preference scores for four patterns across four dimensions. Error bars represent weighted standard deviations. Statistical significance is indicated as ** $p < 0.01$.

the **BASE** ($M_d = 5.5, IQR = 4.0$). On the other hand, Q8 (*I can get useful information through it.*) showed no significant difference ($U = 147, p = 0.13$).

As shown in Fig. 7, participants consistently preferred **active listening** over other conversation patterns in **SIMVIEWS**. Understanding ($M = 3.30, SD = 0.80, p < 0.01$), Engagement ($M = 3.25, SD = 0.97, p < 0.01$), and Satisfaction ($M = 3.35, SD = 0.81, p < 0.01$) both showed significance. And Usefulness ($M = 3.05, SD = 0.99, p = 0.076$) showed no significance.

6.2 Analysis of Conversational Behavior

Two researchers coded and cross-validated the video data, counting each user-agent exchange as one conversational turn; key metrics are summarized in Table 1. Specifically, participants engaged in more conversations in the **SIMVIEWS**, with a higher total number of turns ($M = 8.1, SD = 4.15$) compared to the **BASE** ($M = 3.4, SD = 2.35$). When focusing on the number of participants actively initiated turns (excluding follow-up questions and passive responses), **SIMVIEWS** ($M = 5.1, SD = 3.13$) also outperformed the **BASE** ($M = 2.16, SD = 1.47$). We also examined whether participants engaged in follow-up turns during the study. The number of follow-up questions was higher in the **SIMVIEWS**

($M = 2.6, SD = 2.3$) compared to the **BASE** ($M = 0.96, SD = 1.47$). Additionally, the maximum number of follow-up turns was also higher in the **SIMVIEWS** ($M = 1.4, SD = 0.94$) than in the **BASE** ($M = 0.9, SD = 1.45$), indicating deeper engagement.

Table 1: Turns of Conversation between SIMVIEWS and BASE

Measure	SIMVIEWS	BASE
Total number of turns	8.1 (4.15)	3.4 (2.35)
Participants initiating turns	5.1 (3.13)	2.16 (1.47)
Follow-up turns	2.6 (2.3)	0.96 (1.47)
Max follow-up turns	1.4 (0.94)	0.9 (1.45)

As shown in Table 2, we further analyzed the conversation patterns in the **SIMVIEWS**. Participants engaged in an average of 2.4 **active speaking** ($SD = 1.9$), with 12/20 participants initiating at least one follow-up turn ($M = 1.4, SD = 2.09$), and the maximum number of follow-up turns was 0.9 ($SD = 1.02$). In contrast, during **passive speaking**, 4/20 participants did not respond at all. The overall response rate to agent-initiated prompts was 50%, and only 2 participants engaged in any follow-up turns after their initial reply ($M = 0.25, SD = 0.64$). For agent-to-agent conversations, only 1 participant completely ignores **passive listening**. 15/20 participants chose to join the discussion after listening, with an average of 2.55 **active listening** ($SD = 1.93$), accounting for 21.25% of all listening behaviors. Additionally, 10/20 participants engaged in at least one follow-up turns during **active listening** ($M = 0.94, SD = 1.15$), and a maximum of 0.65 follow-up turns ($SD = 0.75$).

6.3 Comprehension Scores

After participants completed each task, they were asked to take a knowledge test and articulate their understanding of the diverse perspectives presented in the exhibits. The knowledge test included

Table 2: Conversation Pattern Statistics in SIMVIEWS

Measure	Active Speaking	Active Listening	Passive Speaking	Passive Listening
Total numbers	2.4 (1.9)	2.55 (1.93)	0.55 (0.69)	–
Follow-up turns	1.4 (2.09)	0.94 (1.15)	0.25 (0.64)	–
Max follow-up turns	0.9 (1.02)	0.65 (0.75)	0.25 (0.64)	–

10 multiple-choice questions per exhibit, derived from literature and refined by a museum expert. Each question was worth 5 points, with a maximum score of 50 per exhibit. As shown in Table 3. Participants scored 41.5 points ($SD = 9.88$) in the **SIMVIEWS**, compared to an average score of 43.5 points ($SD = 8.44$) in the **BASE**.

For the viewpoint statements, researchers first coded three keywords for each perspective, which were then validated and refined with input from museum experts. Then researchers independently coded participants' statements and scored them as follows: 2 points for a direct keyword match, 1 point for partial relevance, and 0 for no match. Specifically, participants' statements in the **SIMVIEWS** received an average score of 3.21 ($SD = 1.72$), while those in the **BASE** received an average score of 2.43 ($SD = 1.40$).

Table 3: Scores of Knowledge Test and Viewpoint Statement

Measure	SIMVIEWS	BASE
Knowledge Test	41.5 (9.88)	43.5 (8.44)
Viewpoint Statement	3.21 (1.72)	2.43 (1.40)

6.4 Semi-structured Interview

After transcribing all recordings, two researchers reviewed three randomly selected transcripts and applied an open coding approach, combining deductive and inductive techniques to develop a codebook [7]. They resolved disagreements through discussion, and the research team held regular meetings to refine the codebook until a consensus was reached.

6.4.1 Tangible and Credible Presentation of Perspectives. Multiple agents with distinct identities, appearances, and voices made the perspectives more tangible. "Visually presenting three different viewpoints with three distinct tones of voice" (P11). This approach clarified and separated perspectives, making them easier to access. "Having three agents each state their own viewpoint is clearer than having one agent blend all the viewpoints together" (P2). In contrast, participants found the **BASE** more confusing, as all viewpoints came from the same person. "Sometimes, I couldn't tell which viewpoint he was referring to" (P10). Moreover, the single agent reduced the credibility of the perspectives. "The guide was merely paraphrasing others' views, so I felt their statements might not be professional since they weren't experts" (P9). Whereas multi-agents with specialized identities enhanced credibility: "Since they are professionals in their respective fields, I find their words more trustworthy and am more willing to listen" (P13).

6.4.2 Curiosity-driven Proactive Engagement. Moreover, most participants stated that the agents' identities attracted them to engage in the **active speaking** "Their identities make me curious to ask relevant questions" (P19). The identity also provided clear guidance

to participants: "If I have any question, I can ask the agent associated with the perspective" (P3). For example, in the study, P9 actively asked the museum curator: "I noticed that you are a curator. Performance art is usually seen in art exhibitions rather than in museums. so what do you make of this exhibit?". Meanwhile, conflicts between agents also sparked participants' curiosity. "When I hear a conflict, I want to listen and join in" (P9). This increased engagement led to more follow-up questions and extended discussions in the **SIMVIEWS**. "Conflicts make me more eager to participate" (P15). Participants also took clear stances after listening. For instance, P9 tried to persuade the Native American scholar: "He was trying to convince me, which made me want to keep discussing". Data analysis showed that **active listening** was the most favored pattern among participants, with 15/20 choosing to join the conversation while **passive listening**.

6.4.3 Thinking and Memory Enhances Understanding. Conversations enhanced participants' understanding of diverse perspectives. On the one hand, conversations encouraged participants to think deeply, as P18 said, "I tried to put myself in their shoes and later understood their viewpoint". On the other hand, repetition in conversations also reinforced the memory, "He kept repeating the same thing, which deepened my understanding" (P11). In contrast, 6 participants found the **BASE** harder to follow, "He stated viewpoints one after another before I could grasp the previous one" (P1). Results showed that participants in the **SIMVIEWS** ($M = 3.21$) scored higher than those in the **BASE** ($M = 2.43$), suggesting they remembered and understood more diverse perspectives in **SIMVIEWS**.

However, if participants lacked patience and engagement, they might miss key information, which could explain why the **SIMVIEWS** scored slightly lower in some knowledge tests. As P8 noted, "I should have followed up, then I would have known the answer".

7 DISCUSSION

7.1 Presenting Perspectives via Multiple Agents

7.1.1 Clarifying Diverse Perspectives through Multiple Agent and Layered Presentation. Diverse museums present extensive information and varied viewpoints, which can easily overwhelm visitors, especially when multiple perspectives are from a single source. **SIMVIEWS** addresses this challenge by distributing viewpoints across multiple agents, each with a unique professional identity and multimodal features. It allows users to associate each perspective with a specific agent, reducing confusion and enhancing clarity, as P3 noted: "If I have any question, I can ask the agent associated with the perspective". Prior work also supports the effectiveness of multi-agent systems in conveying contrasting views and aiding comprehension [36]. Visual and behavioral differences among agents further reinforce the clarity and memorability of the content [77].

However, some users felt that multi-agent conversations lacked essential background information, creating a barrier to understanding. This may explain why knowledge test scores in **SIMVIEWS** were slightly lower than in **BASE**. They recommended introducing a guide agent to offer basic exhibit information before the perspective-based conversations begin. This layered structure supports users' understanding of perspectives while reducing cognitive load [62],

and aligns with the principles of Learning Hierarchies [75]. Future virtual museums should continue to explore such hierarchical strategies for presenting perspectives clearly and gradually.

7.1.2 Facilitating Proactive Questioning through Characterized Agents. It remains a challenge to present diverse perspectives appropriately while attracting and sustaining a visitor's attention. In **SIMVIEWS**, we assign professional identities to agents to make conversations feel more purposeful and credible, encouraging engagement [36, 66]. Our results also showed that participants in **SIMVIEWS** engaged more actively ($M = 8.10$) than **BASE** ($M = 3.40$). Because participants expressed a greater willingness to interact with characterized visitor agents who appeared more professional than with guide agents who seemed to provide only surface-level information. Participants valued the agents' professional fields (e.g., whether a biologist specialized in tigers or lions) for targeted questioning, rather than detailed identity attributes (e.g., age or name). Research also confirms that expertise boosts agent persuasiveness [50].

Meanwhile, participants expressed a desire for more identity diversity among agents to access richer perspectives [65]. However, some noted that increasing agent variety still couldn't fully capture all perspectives. They suggested tailoring agent recommendations based on user preferences. Prior work has explored using LLMs to personalize museum content or visitor paths [45, 71, 72], pointing to future opportunities for LLM-driven systems to recommend professional perspectives aligned with individual interests.

7.2 Conversations for Understanding and Engagement in Diverse Museums

7.2.1 Deepening Understanding through Discussion and Conversational Feedback. Diverse perspectives increase the complexity of museum content, making it harder to understand. To support comprehension, we designed visitor-to-visitor conversations where users can gain diverse perspectives. Participants reported a deeper understanding through these conversations, with results indicating that **active listening** ($M = 3.30$) was more helpful than **active speaking** ($M = 2.50$) for understanding perspectives. This was because the discussion prompted memory and thinking more effectively than a single-agent interpretation. As P11 noted, "*He repeatedly stated the same point, and this repetition actually deepened my memory.*" Others described how the discussion helped them consider alternative viewpoints, as P18 shared, "*I try to think from his perspective, so later I was able to better understand his viewpoint.*" The viewpoints presented through discussion are expressed and reinforced, thus enhancing the memorability and acceptance of the information.

However, participants noted that the conversational multimodal feedback was limited. They suggested that agents should better simulate physical museum behavior by referencing exhibits through gestures and movement to support understanding [29, 30]. For example, when the ethicist discussed human remains in Lion Attacking a Dromedary, participants expected the agent to walk toward the exhibit and visually direct their attention. Prior work supports this need, showing that synchronized multimodal cues (e.g., text highlights or pointing gestures) enhance comprehension in complex environments [74]. Improving such feedback could further support users' understanding of diverse perspectives.

7.2.2 Enhancing Engagement through Agent Debate and User-Controlled Conversations. Agent debate serves as an effective mechanism for sustaining visitor engagement in virtual museums. Disagreements between visitor agents raised participants' curiosity, fulfilling their desire for "*watching a bustling scene*" and attracting them to actively participate. Some participants noted, "*When I hear a conflict in their discussion, I want to listen, and then I want to join in*". Results also showed that 15/20 participants transitioned from **passive listening** to **active listening** during listening discussions. Such conflicts not only sparked participants' interest but also prompted them to continue questioning and engaging in deeper discussions. We observed that the number of follow-up turns in **SIMVIEWS** ($M = 2.60$) was higher than in **BASE** ($M = 0.96$), and **active listening** had a relatively high number of follow-up turns ($M = 0.94$), suggesting that conflict discussions effectively enhance engagement [51].

In addition, participants suggested leveraging the advantages of virtual museums over physical museums to enhance the user's dominant position [42]. They favored and engaged in **active speaking**, which had the highest number of follow-up turns ($M = 1.4$), while **passive speaking** ($M = 0.25$) had the lowest. Although some expressed concerns about social pressure while speaking, most felt the virtual setting reduced anxiety [35, 79]. Some also hoped to assign a visitor agent to follow them and participate in other conversations together, rather than only joining existing ones. Prior work has explored using natural language in Unity for such control [34], and future research can further expand the conversation pattern and enhance the user's dominant position in virtual museums.

7.3 Limitations and Future Work

Our study has some limitations. First, the exhibits may not have interested all participants, potentially affecting their engagement and overall evaluation. Future research could include a broader range of exhibit themes to attract diverse participants and improve generalizability. Additionally, this study was conducted in a simplified virtual museum without VR devices, which could affect user interaction and engagement [32]. Future studies could explore VR environments to see if multi-agent systems offer a more immersive experience. Lastly, interruptions can effectively enhance the fluency and engagement of conversations. However, our design currently lacks support for interrupting conversations, which could be considered for future implementation. These approaches would lead to a richer, more personalized user experience in virtual museums.

8 CONCLUSION

Diverse perspectives enrich museum experiences, yet most virtual museums still rely on singular, authoritative narratives. Using recent advances in LLMs, we developed SimViews, a multi-agent system that presents diverse perspectives through simulated visitor-to-visitor conversation. Evaluation results show that SimViews effectively supports understanding of diverse perspectives and increases user engagement. This highlights the promise of multi-agent conversation for building more inclusive and diverse virtual museum experiences. We hope our insights will inspire future efforts toward more diverse, equitable, and inclusive virtual museum design.

Acknowledgments

This work is partially supported by Guangzhou Higher Education Teaching Quality and Teaching Reform Project (No. 2024YBJG070), HKUST Practice Research (No. HKUST(GZ)-ROP2025002), Guangzhou-HKUST(GZ) Joint Funding Project (No. 2024A03J0617), Education Bureau of Guangzhou Municipality Funding Project (No. 2024312152), Guangdong Provincial Key Lab of Integrated Communication, Sensing and Computation for Ubiquitous Internet of Things (No. 2023B1212010007), the Project of DEGP (No.2023KCXTD042), and Artificial Intelligence Research and Learning Base of Urban Culture (No. 2023WZJD008).

References

- [1] Rahman Abidin, Nana Supriatna, Didin Saripudin, and Leli Yulifar. 2024. Exploring Critical and Ethnopedagogical Dimensions in Living Museums: Ethnographic Perspectives in Cultural Education Practices. *Journal of Education, Teaching and Learning* 9, 2 (2024), 12–19.
- [2] Amy L Baylor. 2011. The design of motivational agents and avatars. *Educational Technology Research and Development* 59 (2011), 291–300.
- [3] Tessa Beinema, Daniel Davison, Dennis Reidsma, Oresti Banos, Merijn Bruijnes, Brice Donval, Álvaro Fides Valero, Dirk Heylen, Dennis Hofs, Gerwin Huizing, et al. 2021. Agents United: An open platform for multi-agent conversational systems. In *Proceedings of the 21st ACM International Conference on Intelligent Virtual Agents*. 17–24.
- [4] Simon A Black, Amina Fellous, Nobuyuki Yamaguchi, and David L Roberts. 2013. Examining the extinction of the Barbary lion and its implications for felid conservation. *PLoS One* 8, 4 (2013), e60174.
- [5] Jane Blocker. 2001. Failures of Self-Seeing: James Luna Remembers Dino. *PAJ: A Journal of Performance and Art* 23, 1 (2001), 18–32.
- [6] Jane Blocker. 2009. Ambivalent Entertainments: James Luna, Performance, and the Archive. *Grey Room* 37 (2009).
- [7] Virginia Braun and Victoria Clarke. 2012. *Thematic analysis*. American Psychological Association.
- [8] Martin Briesch, Dominik Sobania, and Franz Rothlauf. 2023. Large language models suffer from their own output: An analysis of the self-consuming training loop. *arXiv preprint arXiv:2311.16822* (2023).
- [9] Duncan F Cameron. 1972. The museum: A temple or the forum. *Cahiers d'Histoire Mondiale. Journal of World History. Cuadernos de Historia Mundial* 14, 1 (1972), 189.
- [10] Christine Chin. 2013. Key dimensions of a multicultural art education curriculum. *International Journal of Education & the Arts* 14, 14 (2013).
- [11] Jean Ho Chu and Ali Mazalek. 2019. Embodied engagement with narrative: a design framework for presenting cultural heritage artifacts. *Multimodal Technologies and Interaction* 3, 1 (2019), 1.
- [12] Sheng Kuan Chung. 2009. Presenting cultural artifacts in the art museum: A university-museum collaboration. *Art Education* 62, 3 (2009), 33–39.
- [13] Ralph Cram and Lisa Fletcher. 2014. Picturing the Indian tiger: Imperial Iconography in the Nineteenth century. *Victorian Literature and Culture* 42, 3 (2014), 369–386.
- [14] Costis Dallas. 2004. The presence of visitors in virtual museum exhibitions. In *Numerisation, lien social, lectures colloquium, University of Crete, Rethymnon*. 3–4.
- [15] Rossana Damiano, Tsvi Kuflik, Alan J Wecker, Manuel Striani, Antonio Lieto, Luis Bruni, Nele Kadastik, and Thomas A Pedersen. 2022. Exploring values in museum artifacts in the SPICE project: A preliminary study. In *Adjunct Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization*. 391–396.
- [16] Maira Gatti de Bayser, Melina Alberio Guerra, Paula Cavalin, and Claudio Pinhanez. 2018. Specifying and implementing multi-party conversation rules with finite-state-automata. In *Workshops at the Thirty-Second AAAI Conference on Artificial Intelligence*.
- [17] Maira Gatti de Bayser, Claudio Pinhanez, Heloisa Candello, Marisa Affonso, Mauri Pichiliani Vasconcelos, Melina Alberio Guerra, Paulo Cavalin, and Renan Souza. 2018. Ravel: a mas orchestration platform for human-chatbots conversations. In *The 6th International Workshop on Engineering Multi-Agent Systems (EMAS@ AAMAS 2018), Stockholm, Sweden*.
- [18] Rahul R Divekar, Hui Su, Jeffrey O Kephart, Maira Gratti DeBayser, Melina Guerra, Xiangyang Mou, Matthew Peveler, and Lisha Chen. 2020. HUMAINE: human multi-agent immersive negotiation competition. In *Extended abstracts of the 2020 CHI conference on human factors in computing systems*. 1–10.
- [19] Yilun Du, Shuang Li, Antonio Torralba, Joshua B Tenenbaum, and Igor Mordatch. 2023. Improving factuality and reasoning in language models through multiagent debate. *arXiv preprint arXiv:2305.14325* (2023).
- [20] Marta Dynel. 2011. Revisiting Goffman's postulates on participant statuses in verbal interaction. *Language and Linguistics Compass* 5, 7 (2011), 454–465.
- [21] Debora Engelmann, Alisson Panisson, Renata Vieira, Jomi Hubner, Viviana Mascalchi, and Rafael Bordini. 2023. MAIDS-a Framework for the Development of Multi-Agent Intentional Dialogue Systems. (2023).
- [22] John Falk. 2016. Museum audiences: A visitor-centered perspective. *Loisir et Société/Society and Leisure* 39, 3 (2016), 357–370.
- [23] Amina Fellous-Djardini, Nobuyuki Yamaguchi, Hadi Al Hikmani, Brahim Hadjane, and Simon A Black. 2023. The natural history of lions in North Africa and the relevance of their depiction in the "Lions Attacking a Dromedary" diorama. *Curator: The Museum Journal* 66, 3 (2023), 419–439.
- [24] Erving Goffman. 1981. *Forms of talk*. University of Pennsylvania Press.
- [25] Anne-Britt Gran, Nina Lager Vestberg, Peter Booth, and Anne Ogundipe. 2019. A digital museum's contribution to diversity—a user study. *Museum management and curatorship* 34, 1 (2019), 58–78.
- [26] Susie Green. 2006. *Tiger*. Reaktion Books.
- [27] Taicheng Guo, Xiuying Chen, Yaqi Wang, Ruidi Chang, Shichao Pei, Nitesh V Chawla, Olaf Wiest, and Xiangliang Zhang. 2024. Large language model based multi-agents: A survey of progress and challenges. *arXiv preprint arXiv:2402.01680* (2024).
- [28] Elizabeth S Hawley. 2016. James Luna and the paradoxically present vanishing Indian. *Contemporaneity: Historical Presence in Visual Culture* 5 (2016), 5–26.
- [29] Hoang Phuoc Ho, Vani Ramesh, Ivo Zaloudek, Delaram Jaydani Rikhtehgar, and Shenghui Wang. 2025. Enhancing Visitor Engagement in Interactive Art Exhibitions with Visual-Enhanced Conversational Agents. In *Proceedings of the 30th International Conference on Intelligent User Interfaces*. 660–671.
- [30] Eva Hornecker and Luigina Ciolfi. 2022. *Human-computer interactions in museums*. Springer Nature.
- [31] Han-Yin Huang and Cynthia CS Liem. 2022. Social inclusion in curated contexts: Insights from museum practices. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*. 300–309.
- [32] Wolfgang Hürst, Ferdinand de Coninck, and Xhi Jia Tan. 2016. Complementing artworks to create immersive VR museum experiences. In *Proceedings of the 13th international conference on advances in computer entertainment technology*. 1–6.
- [33] International Council for Museums. 2021. ICOM Announces the Alternative Museum Definition That Will Be Subject to Vote. <https://icom.museum/en/news/icom-announces-the-alternative-museum-definition-that-will-be-subject-to-a-vote/>. Retrieved April 30, 2022.
- [34] Ray Ito and Junichiro Takahashi. 2024. Demo Paper: A Game Agents Battle Driven by Free-Form Text Commands Using Code-Generation LLM. *arXiv preprint arXiv:2405.11835* (2024).
- [35] Isabel L Kampmann, Paul MG Emmelkamp, Dwi Hartanto, Willem-Paul Brinkman, Bonne JH Zijlstra, and Nekhmedin Morina. 2016. Exposure to virtual social interactions in the treatment of social anxiety disorder: A randomized controlled trial. *Behaviour research and therapy* 77 (2016), 147–156.
- [36] Reshmaresh B Kantharaju, Dominik De Franco, Alison Pease, and Catherine Pelachaud. 2018. Is two better than one? Effects of multiple agents on user persuasion. In *Proceedings of the 18th international conference on intelligent virtual agents*. 255–262.
- [37] Ivan Karp. 1991. *Exhibiting cultures: The poetics and politics of museum display*. Smithsonian Institution.
- [38] Gaynor Kavanagh. 2005. *Museum provision and professionalism*. Routledge.
- [39] Claudette Lauzon. 2013. 'Playing Indian': Art, aboriginality, and the politics of identification. *Performance Research* 18, 1 (2013), 158–161.
- [40] Steven D Levine and Ivan Karp. 1991. *Exhibiting cultures: The poetics and politics of museum display*. Smithsonian Institution Press.
- [41] Guohao Li, Hasan Hammoud, Hani Itani, Dmitrii Khizbulin, and Bernard Ghanem. 2023. Camel: Communicative agents for "mind" exploration of large language model society. *Advances in Neural Information Processing Systems* 36 (2023), 51991–52008.
- [42] Yue Li, Paul Tennent, and Sue Cobb. 2019. Appropriate control methods for mobile virtual exhibitions. In *VR Technologies in Cultural Heritage: First International Conference, VRTCH 2018, Brasov, Romania, May 29–30, 2018, Revised Selected Papers 1*. Springer. 165–183.
- [43] Tian Liang, Zhiwei He, Wenxiang Jiao, Xing Wang, Yan Wang, Rui Wang, Yujiu Yang, Zhaopeng Tu, and Shuming Shi. 2023. Encouraging divergent thinking in large language models through multi-agent debate. *arXiv preprint arXiv:2305.19118* (2023).
- [44] Weijane Lin, Wen-Ting Lo, and Hsiu-Ping Yueh. 2019. How the multimodal media in augmented reality affects museum learning Experience. In *2019 12th Asia Pacific Workshop on Mixed and Augmented Reality (APMAR)*. IEEE, 1–4.
- [45] Chao Liu, Chi San Cheung, Mingqing Xu, Zhongyue Zhang, Mingyang Su, and Mingming Fan. 2024. Toward Facilitating Search in VR With the Assistance of Vision Large Language Models. In *Proceedings of the 30th ACM Symposium on Virtual Reality Software and Technology*. 1–14.
- [46] Chao Liu, Mingyang Su, Yan Xiang, Yuru Huang, Yiqian Yang, Kang Zhang, and Mingming Fan. 2025. Toward Enabling Natural Conversation with Older Adults via the Design of LLM-Powered Voice Agents that Support Interruptions and Backchannels. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–22.

- [47] Pengfei Liu, Weizhe Yuan, Jinlan Fu, Zhengbao Jiang, Hiroaki Hayashi, and Graham Neubig. 2023. Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *Comput. Surveys* 55, 9 (2023), 1–35.
- [48] Yafei Liu, Hongjin Qian, Hengpeng Xu, and Jinmao Wei. 2020. Speaker or listener? the role of a dialog agent. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, 4861–4869.
- [49] Ziyi Liu, Zhengzhe Zhu, Lijun Zhu, Enze Jiang, Xiyun Hu, Kylie A Pepple, and Karthik Ramani. 2024. ClassMeta: Designing Interactive Virtual Classmate to Promote VR Classroom Participation. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–17.
- [50] Amama Mahmood and Chien-Ming Huang. 2022. Effects of rhetorical strategies and skin tones on agent persuasiveness in assisted decision-making. In *Proceedings of the 22nd ACM International Conference on Intelligent Virtual Agents*. 1–8.
- [51] Irena Mišurac-Zorica and Maja Cindrić. 2012. Advantages of discussion and cognitive conflict as an instructional method in contemporary teaching of mathematics. *Zbornik Instituta za pedagoška istraživanja* 44, 1 (2012), 92–110.
- [52] Salwa Mikdadi Nashashibi. 2003. Visitor voices in art museums: The visitor-written label. *Journal of Museum Education* 28, 3 (2003), 21–25.
- [53] Barry Natusch. 2015. The role of Technology in Participatory Museum experiences. *Nihon University Research Bulletin* 77 (2015), 79–99.
- [54] André Neves, Silvio Meira, Filipe Calegário, Rui Belfort, and Marcello Bressan. 2024. squad.ai: A Multi-agent System Built on LLMs, Incorporating Specialized Embeddings and Sociocultural Diversity. In *International Conference on Human-Computer Interaction*. Springer, 86–96.
- [55] Miranda AM Niittynen. 2023. Lions Attacking a Dromedary: The Verreaux brothers, imperial taxidermy, and postmortem bodily rights. *Curator: The Museum Journal* 66, 3 (2023), 441–457.
- [56] Mark O'Neill. 2008. Museums, professionalism and democracy. *Cultural trends* 17, 4 (2008), 289–307.
- [57] Jeongeon Park, Bryan Min, Xiaojuan Ma, and Juho Kim. 2023. Choicemates: Supporting unfamiliar online decision-making with multi-agent conversational interactions. *arXiv preprint arXiv:2310.01331* (2023).
- [58] Yannik Porsché and Yannik Porsché. 2018. Public Interactions on the Museum Stages. *Public Representations of Immigrants in Museums: Exhibition and Exposure in France and Germany* (2018), 223–319.
- [59] Chen Qian, Xin Cong, Cheng Yang, Weize Chen, Yusheng Su, Juyuan Xu, Zhiyuan Liu, and Maosong Sun. 2023. Communicative agents for software development. *arXiv preprint arXiv:2307.07924* (2023).
- [60] George E Raptis, Giannis Kavvetsos, and Christina Katsini. 2021. MuMIA: Multimodal interactions to better understand art contexts. *Applied Sciences* 11, 6 (2021), 2695.
- [61] Rebecca Reynolds. 2011. Reinventing the forum: Multiple perspectives, information transmission and new technology. *Museum Management and Curatorship* 26, 1 (2011), 45–62.
- [62] Rufat Rzayev, Gürkan Karaman, Niels Henze, and Valentin Schwind. 2019. Fostering virtual guide in exhibitions. In *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services*. 1–6.
- [63] Christiane Oved Særkjær. 2024. Visitor voices in the museum space: sharing art experiences as a dialogic mode in museum communication. *Conjunctions* 11, 1 (2024), 1–13.
- [64] Briane Paul V Samson and Yasuyuki Sumi. 2020. Are two heads better than one? exploring two-party conversations for car navigation voice guidance. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–9.
- [65] Katina Sawyer, Nicholas Salter, and Christian Thoroughgood. 2013. Studying individual identities is good, but examining intersectionality is better. *Industrial and Organizational Psychology* 6, 1 (2013), 80–84.
- [66] William H Sharp and Marc M Sebrechts. 2020. Impact of Perceived Agent Expertise on Trust in Computer Agent Recommendations. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 64. SAGE Publications Sage CA: Los Angeles, CA, 1355–1359.
- [67] Christine E Sleeter and Carl A Grant. 2008. *Making choices for multicultural education: Five approaches to race, class and gender*. John Wiley & Sons.
- [68] Ramesh Srinivasan, Katherine M Becvar, Robin Boast, and Jim Enoe. 2010. Diverse knowledges and contact zones within the digital museum. *Science, technology, & human values* 35, 5 (2010), 735–768.
- [69] Xiangru Tang, Anni Zou, Zuoosheng Zhang, Yilun Zhao, Xingyao Zhang, Arman Cohan, and Mark Gerstein. 2023. Medagents: Large language models as collaborators for zero-shot medical reasoning. *arXiv preprint arXiv:2311.10537* (2023).
- [70] David Traum and Jeff Rickel. 2002. Embodied agents for multi-party dialogue in immersive virtual worlds. In *Proceedings of the first international joint conference on Autonomous agents and multiagent systems: part 2*. 766–773.
- [71] Georgios Trichopoulos. 2023. Large language models for cultural heritage. In *Proceedings of the 2nd International Conference of the ACM Greek SIGCHI Chapter*. 1–5.
- [72] Georgios Trichopoulos, Markos Konstantakis, Georgios Alexandridis, and George Caridakis. 2023. Large language models as recommendation Systems in Museums. *Electronics* 12, 18 (2023), 3829.
- [73] Zhiyang Wang, Jina Lee, and Stacy Marsella. 2013. Multi-party, multi-role comprehensive listening behavior. *Autonomous Agents and Multi-Agent Systems* 27 (2013), 218–234.
- [74] Zhan Wang, Lin-Ping Yuan, Liangwei Wang, Bingchuan Jiang, and Wei Zeng. 2024. Virtuawander: Enhancing multi-modal interaction for virtual tour guidance through large language models. In *Proceedings of the CHI conference on human factors in computing systems*. 1–20.
- [75] Richard T White. 1973. Research into learning hierarchies. *Review of Educational Research* 43, 3 (1973), 361–375.
- [76] Benfeng Xu, An Yang, Junyang Lin, Quan Wang, Chang Zhou, Yongdong Zhang, and Zhendong Mao. 2023. Expertprompting: Instructing large language models to be distinguished experts. *arXiv preprint arXiv:2305.14688* (2023).
- [77] Philipp Yasrebi-Soppa, Jobst-Julius Bartels, Sebastian Viehaus, Pascal Reuss, and Klaus-Dieter Althoff. 2020. Visualizing the behavior of CBR agents in an FPS Scenario. (2020).
- [78] Bangguo Yu, Hamidreza Kasaei, and Ming Cao. 2023. Co-NavGPT: Multi-robot cooperative visual semantic navigation using large language models. *arXiv preprint arXiv:2310.07937* (2023).
- [79] Samaneh Zamanifar and Andrew Robb. 2023. Social virtual reality is my therapist: overcoming social anxiety disorder through using social virtual reality. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–6.
- [80] Yu Zhang, Jingwei Sun, Li Feng, Cen Yao, Mingming Fan, Liuxin Zhang, Qianying Wang, Xin Geng, and Yong Rui. 2024. See Widely, Think Wisely: Toward Designing a Generative Multi-agent System to Burst Filter Bubbles. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–24.
- [81] Zhiling Zheng, Oufan Zhang, Ha L Nguyen, Nakul Rampal, Ali H Alawadhi, Zichao Rong, Teresa Head-Gordon, Christian Borgs, Jennifer T Chayes, and Omar M Yaghi. 2023. Chatgpt research group for optimizing the crystallinity of mofs and cofs. *ACS Central Science* 9, 11 (2023), 2161–2170.