

Effects of Embodiment on Generic and Content-Specific Intelligent Virtual Agents as Exhibition Guides

Susanne Schmidt¹ , Gerd Bruder²  and Frank Steinicke¹ 

¹Universität Hamburg, Germany
²University of Central Florida, USA

Abstract

Intelligent Virtual Agents (IVAs) received enormous attention in recent years due to significant improvements in voice communication technologies and the convergence of different research fields such as Machine Learning, Internet of Things, and Virtual Reality (VR). Interactive conversational IVAs can appear in different forms such as voice-only or with embodied audio-visual representations showing, for example, human-like contextually related or generic three-dimensional bodies. In this paper, we analyzed the benefits of different forms of virtual agents in the context of a VR exhibition space. Our results suggest positive evidence showing large benefits of both embodied and thematically related audio-visual representations of IVAs. We discuss implications and suggestions for content developers to design believable virtual agents in the context of such installations.

CCS Concepts

• **Human-centered computing** → **Mixed / augmented reality; Virtual reality; Empirical studies in HCI;**

1. Introduction

Inspired by science fiction media, such as the movies *Her* (2013) and *Blade Runner* (2017)—stories that show the potential of disembodied or embodied Intelligent Virtual Agents (IVAs) integrated into our daily social life—we have seen a large public interest in related technologies.

Over the last years, voice-controlled agents were embedded in consumer devices such as Amazon’s Echo or Apple’s HomePod, and connected to home appliances to provide an intuitive and natural form of interaction with smart home environments and as a means to access information from the Internet or to control smart home devices [KBB*18]. Beyond home uses, smart services provided by IVAs are popular as they can be accessed through ubiquitous smartphone technologies and can be implemented for professional applications such as in the form of educational audio guides in museums or audio-visual presentations for mixed media installations or exhibits. In particular, in situations where the demands for individual support or care exceed the supply of specialized trained personnel, such as museum guides, training scenarios, or rehabilitation, these IVAs are a promising solution that can complement human professionals [NKH*18]. With the current convergence of different research fields such as Machine Learning, Internet of Things, and Virtual Reality (VR), it seems reasonable to assume that people will be confronted with an increasing amount of such services, which poses new challenges to the interface designers, particularly in terms of social interaction and integration.

While most of the currently used services are limited to audio or flat 2D visual representations, VR and Augmented Reality (AR) technology can add a new dimension by providing a spatially registered 3D virtual body to complement the voice. Human-like VR/AR representations can enrich the communicative channels that convey the agent’s status and intentions to interlocutors with gestures and other forms of social behaviors. Moreover, they can be registered spatially with their environment, which enables a more direct form of spatial interaction compared to voice-only interaction. This is particularly interesting in situations that have a strong spatial component such as art installations and museum exhibitions, since spatial relations are usually harder to communicate via speech than with gestures [Ali05]. Therefore, it may be beneficial to provide an IVA with a virtual body, which could also increase the user’s feeling of co-presence, i. e., raising the visitor’s sense of being together with the content on display. For museum exhibits this could be strengthened, for instance, by choosing a historical figure as the agent’s representation. Through the encounter with a contemporary witness, visitors get to know the subject matter from a personal perspective, which may increase interest in the historical events as well as empathy with the people involved.

In this paper, we present a human-subject study that we performed in a historical exhibition context to understand the importance of embodiment and thematic closeness of the visual representation of virtual agents. Therefore, we perform a case study and analyzed the effects of different intelligent virtual museum guides in the scope of an exhibition related to the Apollo 11 mission on the

elicited sense of co-presence, social presence, and effectiveness in communicating a sense of social competence and trust. The study was designed to evaluate the following two research questions:

1. Do embodied virtual guides perform significantly better than voice-only guides in terms of co-presence, social presence, credibility, and the ability to impart knowledge?
2. Do thematically close content-related guides perform better than generic guides in terms of the above mentioned metrics?

Our results suggest a positive evidence showing large benefits for both embodied and thematically close audio-visual representations. We discuss implications and suggestions for user interface and content developers to design believable virtual agents in the context of such installations.

The remainder of this paper is structured as follows. Section 2 gives an overview of related work on IVAs in VR/AR, embodiment, and presence. In Section 3 we present the conducted user study and discuss the results. Section 4 concludes the paper and discusses future research.

2. Related Work

In this section, we resume work related to agent embodiment as well as presence and confidence in intelligent virtual agents.

Intelligent Virtual Agents in VR/AR. Different forms of IVAs were proposed and evaluated throughout Milgram's reality-virtuality continuum, which were surveyed by Holz et al. [HCO*11, HDO09] and Norouzi et al. [NKH*18], showing the potential of VR/AR agents, but also challenges related to creating a high sense of social interaction and connection between users and virtual agents. For instance, Obaid et al. [ONP11, ODK*12] showed that the physiological arousal of users in VR/AR depends on an agent's behavior associated with cultural differences, e.g., related to gaze behavior and interpersonal distances. Lee et al. [LBHW18] found that the proxemics during interaction with IVAs in AR differs significantly from those between real humans, with users giving virtual agents more space than they would a real person. Kim et al. [KMB*17, KBW17] showed that visual conflicts in AR such as occlusion and dual occupancy between virtual agents and physical objects can significantly impair their social connection with users. However, despite the challenges related to realistic and/or effective social interaction, a large number of applications could benefit from IVAs [NKH*18, KBB*18]. For further information on IVAs, we refer to Magnenat-Thalmann et al. [MTPC08], which provides a literature review of promising application fields for IVAs including interactive virtual guides in cultural heritage sites, museums, art installations, and related fields.

Embodiment. A large body of literature focused on the question if and how virtual agents should be embodied for effective interaction with real humans. Dehn and van Mulken [Dv00] presented a literature review on the embodiment of virtual agents, showing that the early prototypes of embodied agents in the last millennium had mixed effects on human-agent interaction, which sometimes would improve agent-based user interfaces while often they would not provide any benefits. A newer literature meta review by Yee et al. [YBR07] showed benefits for virtual agents with a

face over those implemented just by voice or via text on a computer screen. Moreover, they found that showing a face in general is more important than the realism of the visual presentation or the behavior of the agent. Even an abstract face of a virtual agent can provide important social cues for human-agent interaction, such as communicating visual attention due to the gaze direction of the eyes. Over the last decade, a large number of studies were conducted using VR technologies, which documented the psychological benefits of agent embodiment based on holographic or stereoscopic 3D displays, with recent work showing benefits for rapport as well as realistic social interaction [WLK*15, KNB*17]. Demeur et al. [DNP11] showed that virtual agents in social situations can appear more believable and elicit a higher sense of competence and warmth if they are embodied. Recent studies by Kim et al. [KBW17, KNB*17, KMB*17] indicate that it is more challenging for embodied virtual agents in AR to appear realistic and plausible to users. This is due to the fact that they are more affected by differences in visual appearance between their body and real-world objects as well as the fact that they generally have less control over their environment than is the case in VR, which can limit how virtual agents in AR are perceived and utilized.

Social and Co-Presence. A generalizable metric for the effectiveness of IVAs in VR/AR is their ability to convey an illusion of being perceived as a real social entity sharing the same space with a real person, called *social presence* and *co-presence*. Co-presence denotes the sense of "being together" and social presence the sense of "being socially connected" [HB04]. Blascovich et al. define social presence as "the degree to which one believes that he or she is in the presence of, and dynamically interacting with, other veritable human beings" [Bla02, BLB*02]. In general, the sense of "being there" is denoted *presence*, which can be further refined with the concepts of *place illusion* and *plausibility illusion* introduced by Slater [Sla09]. Plausibility illusion indicates that "the scenario being depicted is actually occurring" with a "credible scenario and plausible interactions between the participant and objects and virtual characters in the environment." Various studies were conducted on IVAs in VR/AR aimed at identifying effects of IVA characteristics on the sense of social and co-presence during interaction using measures such as questionnaires, physiological responses, and behavioral differences, e.g., related to proxemics [FAJ*14]. For instance, Lee et al. [LBHW18] found that more realistic multimodal feedback related to footstep vibrations of virtual agents that are transmitted through the floor using subwoofer devices could significantly improve subjective ratings of social and co-presence in AR. Moreover, they found that the limited field of view of current-state optical see-through AR head-mounted displays (HMDs) can negatively impact social and co-presence and cause less natural proxemic behavior near a virtual agent. Nowak and Biocca [NB03] evaluated different types of IVAs and found to their surprise that a higher anthropomorphism of IVAs reduced the sense of social and co-presence, which they explained stating that higher anthropomorphism might reinforce a person's expectations about realistic behaviors of the virtual agent, which were not met in their study. Chuah et al. [CRW*13] proposed hybrid IVAs with partially physical body parts (legs) for medical applications, suggesting that a higher physicality of IVAs could encourage higher social presence. This result is similar to what Kim et al. [KNB*17] found for robotic

IVAs. Kim et al. [KBMW16] further observed that the sense of social and co-presence depends on the personality of the real interlocutors, showing that extroverted participants reached a higher social presence with IVAs than introverted participants.

3. User Study

In this section, we describe the user study that we conducted to understand the effects of a virtual agent's embodiment relative to the thematic context on the example of a museum exhibition. In our simulated case study, we explore an exhibition, which simulated four episodes of the first manned moon landing. Each episode was presented by a different virtual guide in randomized order: (i) a generic virtual character or (ii) a thematically close content-related astronaut character, each presented either as (iii) a disembodied voice (as known from voice-controlled agents such as Amazon's Echo) or (iv) a stereoscopic 3D embodied representation.

3.1. Participants

In total, 24 participants (17 male and 7 female; ages from 19 to 39, $M = 25.1$) participated in our experiment. All of them were students or staff members of the local Department of Engineering and Computer Science. None of the participants reported any visual or motor impairments that could affect the results of our experiment.

3.2. Material

The experiment was conducted in a four-sided CAVE-like environment with three walls and the floor as projection surfaces. The floor area of the environment was 4.20 by 3.15 meters. Three projectors of type Optoma EH320USTi and one Optoma GT1080e were used to project rendered virtual imagery, each providing a resolution of 1920×1080 at a refresh rate of 120 Hz. All projectors supported synchronized stereoscopic display. Participants wore shutter glasses of type Optoma ZF2300. A six degrees of freedom rigid body target with passive markers was attached to the shutter glasses to track their position and orientation within the environment at 60 Hz using a 7-camera ARTTRACK2 optical tracking system. The voice of the virtual guides was presented to participants via noise-canceling headphones of type Bose QuietComfort 25, with a compatible Bluetooth receiver to make them wireless. Hence, participants were not restricted in their movement and were able to walk around virtual objects in the CAVE freely. Figure 1a shows the experimental setup.

For our case study, we presented four episodes of the Apollo 11 mission, for which we used different models of a scaled-down Saturn V rocket with a launch pad, the interior of the Columbia command module, a scale model of the lunar module, and the moon surface with the American flag as well as scientific experiments (see Figures 1c to 1f). All of the shown virtual models are of historical relevance and both originals and physical replicas are currently on display at museums across the U.S., including the National Air and Space Museum and the Kennedy Space Center. If available, original footage such as a 3D scan of the command module and NASA photographs of the lunar surface was used in order to build detailed models.

We created four versions of the IVA used in the experiment (see Figure 1b):

- **EC:** The Embodied thematically Close character was modeled as an astronaut with a space suit. The astronaut's face was generated using original footage of Neil Armstrong.
- **EG:** The Embodied more Generic virtual character was designed to match a caucasian museum guide wearing a shirt and dress pants. In order to prevent any preference towards one of the guides due to sympathy, we used similar basic facial characteristics for the civilian guide. However, variations of the textures, facial hair and general hair style were made to ensure that the civilian and the astronaut were not perceived as the same person.
- **DC:** The Disembodied voice of the thematically Close astronaut character was identical to that condition except for the visual feedback of the agent.
- **DG:** The Disembodied voice of the more Generic character matched the embodied condition except for the visual feedback.

To increase the level of realism, we added idle behaviors to the embodied virtual guides, and they made eye contact with the user as a real guide would do in a one-on-one conversation.

We anticipated that it would be important that the types of virtual agents would be perceived as different characters and not as the same character dressed in a different way. Hence, we performed a small survey with 20 respondents before the study to fine-tune and validate this aspect of the study.

In the embodied conditions, the agent's lip movements were matched with the spoken text via the Oculus Lip Sync plug-in. The audio track of the guides was created with the Oddcast Vocalware text-to-speech engine. For the astronaut, additional post-processing in Audacity was applied to simulate the sound of radio transmissions at that time.

The four episodes of the Apollo 11 mission provided educational information to the participants, narrated by the virtual guides. The assignment of a virtual guide to the four episodes was randomized. The educational content differed between the four episodes, but it was the same for all guides, except for the narrative point of view: The thematically close astronaut told the story from a first-person perspective and called "his" companions by their given names, while the more generic museum guide told the story from a third-person perspective. The four episodes included the following content (see Figures 1c to 1f):

- **Episode 1:** The preparation of the Apollo 11 mission and its launch at the Kennedy Space Center as well as technical details on the Saturn V rocket.
- **Episode 2:** The three-day journey of the crew to the lunar orbit inside the Command Module with a focus on the roles of Armstrong and Collins.
- **Episode 3:** The descent of Aldrin and Armstrong to the lunar surface using the lunar module as well as the first steps of a man on the moon.
- **Episode 4:** The duties of the astronauts at the landing site, including the flag planting and scientific experiments.

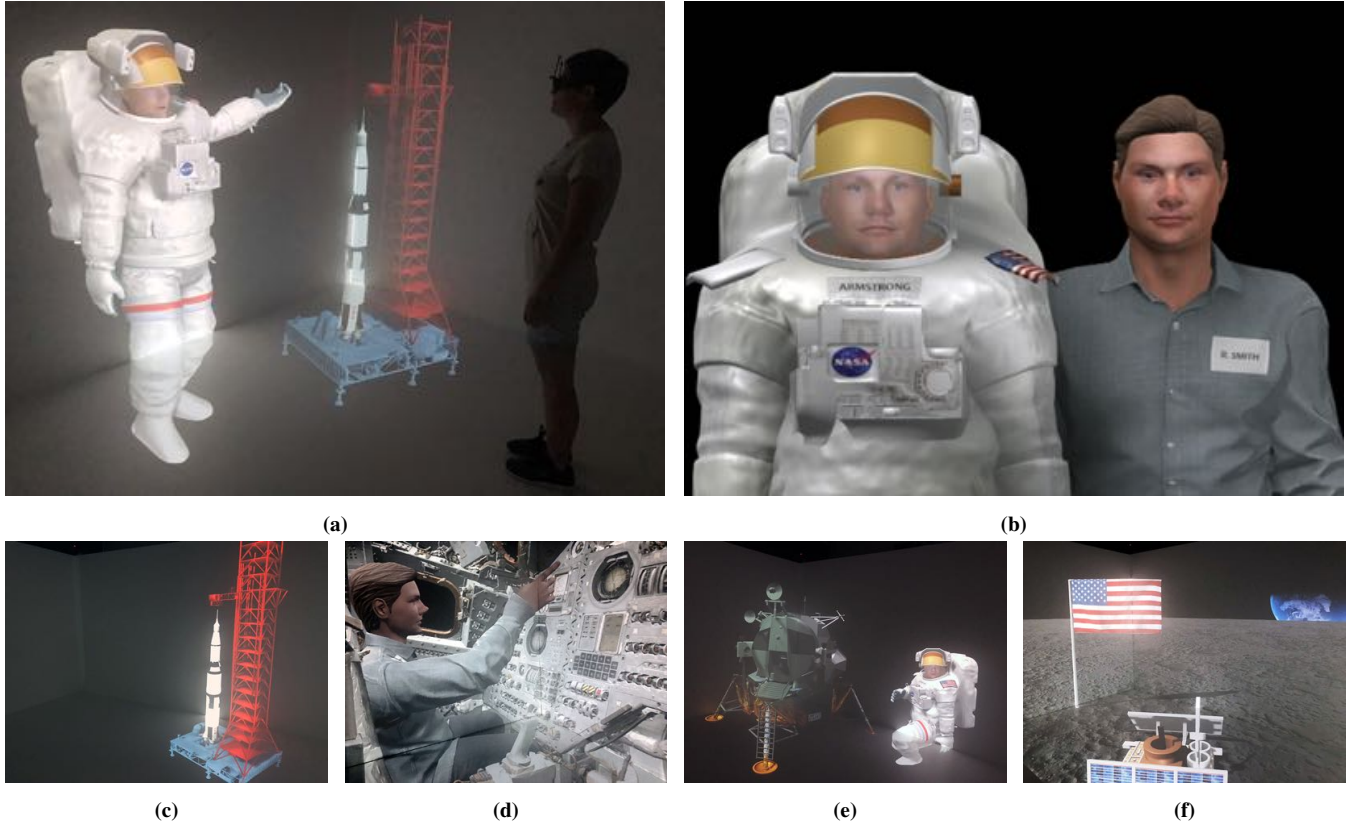


Figure 1: (a) Photo showing the experimental setup, (b) the two guides in their embodied version, and (c) - (f) photos of the four episodes with exemplary guides.

3.3. Methods

We used a within-subjects design using two factors with two levels each: *agent embodiment* (embodied vs. disembodied) and *thematic closeness* (astronaut vs. museum guide). Each participant experienced all four episodes and all four agents EC, EG, DC, and DG described above in randomized order.

Prior to the study, each participant completed a consent form and a demographics questionnaire. Afterwards, participants were guided into the CAVE-like environment by following a virtual 3D floating globe. Participants were introduced to the display technology, had time to familiarize themselves with the system and the stereoscopic display, and then they were informed about the context of the study and the Apollo 11 mission scenario.

After this introductory phase, the main study started with the first of the four episodes of the Apollo 11 mission. Each episode took around three minutes to complete. Participants were allowed to move about the space in the experimental room freely. During the episodes, one of the four guides was present and gave a presentation on the virtual space models on exhibition in the CAVE.

After each episode, the participants were asked to rate their experience using subscales of the Temple Presence Inventory [LDW09] as well as questionnaires that address the agent's credibility and the subjective knowledge gain.

We further ran participants through an “exam” on the presented educational content of the episode they just experienced, assessing how much of the information they actively perceived and could remember. The exam was chosen as a meaningful measure of the guides' quality, since museums usually have an educational mandate. While the visitor is not expected to learn all facts that are presented within an exhibition, the ability to provide interesting information that stick in the visitors' minds is of great value to any public educational institution. In this sense, the exam should give an idea on how successful a guide was to tell a memorable story rather than providing a generalizable percentage of learned facts. Initially, we planned for the exam to be completed without prior notice of the participants at the end of the study. However, a pre-study with ten participants revealed that only a minority of the users paid attention to any of the spoken text and the majority understood it more as an educational entertainment experience. We therefore decided to announce the exam before the study. For each episode a set of 12 questions was prepared, which were similar in terms of their memorizability. They were grouped into four categories: numerical, spatial, social, and visual facts. Numerical questions included sizes, weights, quantities, and periods of time. In spatial tasks, participants had to point at a specific location within a picture of the according scene. This location was described during the episode and was usually supported by a gesture in the embodied conditions. So-

cial facts referred to stories that were experienced by the crew and members of the mission. Visual features were not mentioned by the guide, but could be observed in the presented scene. The exam was conducted orally to ensure that responses, which were guessed or already known before the study, could be identified.

After the exam was finished, participants were guided to the next episode and all steps were repeated. The second episode differed from the other scenes since participants were seated in the center of the CAVE. At the end of all episodes, participants were confronted with all four guides for a second time and had to compare them in an additional questionnaire. The entire study took around 45 to 60 minutes per participant.

3.4. Results

We evaluated the effect of the two factors *agent embodiment* and *thematic closeness* on several subjective and objective measures using multiple two-way repeated measures ANOVAs. The normality assumption was not met in a few cases, however, the ANOVA tolerates moderate deviations from normality, as was shown in several studies [GPS72, HRH092, LKK96].

3.4.1. Presence

Different aspects of presence were measured using the Temple Presence Inventory (TPI) [LDW09]. We focused on four dimensions of the TPI: spatial presence, active social presence, presence as social actor, and presence as engagement. Each dimension involved three to seven items that were measured on a 7-point Likert scale. We ran a two-way repeated measures ANOVA that revealed a significant main effect of agent embodiment on spatial presence ($F(1, 23) = 25.822, p < 0.001, \eta_p^2 = 0.529$), active social presence ($F(1, 23) = 12.181, p = 0.002, \eta_p^2 = 0.346$), presence as social actor ($F(1, 23) = 299.404, p < 0.001, \eta_p^2 = 0.929$), and presence as engagement ($F(1, 23) = 7.516, p = 0.012, \eta_p^2 = 0.246$). Thematic closeness only showed one significant main effect on presence as social actor ($F(1, 23) = 4.420, p = 0.047, \eta_p^2 = 0.161$). No other main effect or interaction effect was significant. The results of the TPI are illustrated in Figure 2a.

3.4.2. Learning

Scores of the oral exam were added up per participant and category, with a score of 3 corresponding to the maximum value of 100%. The results of one participant had to be removed from the data, because he admitted to know several of the tested facts even without the guides due to prior knowledge on the moon landing. The remaining scores were pooled according to the four categories as illustrated in Figure 2b. An ANOVA revealed a significant main effect of agent embodiment on the test scores in the category of visual facts ($F(1, 22) = 8.933, p = 0.007, \eta_p^2 = 0.289$). Apart from this, no other effects on the learning results could be found.

In addition to the objective exam, we also wanted to learn more about the subjective impression of the participants regarding their knowledge gain through the guided presentations. After each episode, before the oral exam, we asked them to make a rough estimate on how many facts they are still able to recall now and in

one week. We ran another ANOVA and found a significant main effect of embodiment on the perceived number of long-term memorized facts ($F(1, 23) = 16.403, p < 0.001, \eta_p^2 = 0.416$), but not on the number of short-term memorized facts ($F(1, 23) = 3.185, p = 0.088, \eta_p^2 = 0.122$).

3.4.3. Credibility

For evaluation of the credibility of guides, we used a scale introduced by McGloin et al. [MNW14]. For each participant, an overall score was constructed from five questionnaire responses to the following bipolar adjective items: "unintelligent to intelligent", "uninformed to informed", "unreliable to reliable", "incompetent to competent", and "untrustworthy to trustworthy". We analyzed the results with a two-way repeated measures ANOVA. Analysis revealed a main effect of agent embodiment on credibility ($F(1, 23) = 5.842, p = 0.024, \eta_p^2 = 0.203$), indicating a significant difference between embodied guides ($M = 5.550, SD = 0.888$), and guides with voice only ($M = 5.254, SD = 1.030$).

3.4.4. User Experience

Besides the aforementioned influence of agent embodiment and thematic closeness on perceived presence, agent credibility and learning, we were also interested in the general experience of users while interacting with the guides. For this purpose, we measured six dimensions of user experience with the UEQ [LHS08]. Participants of the study were asked to provide ratings on 26 items using a 7-point Likert scale. We stressed the point that all responses should be based on the impression of the guide only, without including the virtual scene. This is because the virtual objects were only used in the context of the present study and are no inherent part of applications with AI agents in general. For example, a museum could also incorporate a virtual guide to present real physical exhibits instead of virtual ones. We ran two-way repeated measures ANOVAs for the six dimensions of UEQ. We found a significant main effect of agent embodiment on attractiveness ($F(1, 23) = 8.837, p = 0.007, \eta_p^2 = 0.278$), perspicuity ($F(1, 23) = 8.307, p = 0.008, \eta_p^2 = 0.265$), stimulation ($F(1, 23) = 26.527, p < 0.001, \eta_p^2 = 0.536$), and novelty ($F(1, 23) = 82.786, p < 0.001, \eta_p^2 = 0.783$). Thematic closeness also showed a main effect on attractiveness ($F(1, 23) = 7.212, p = 0.013, \eta_p^2 = 0.239$), stimulation ($F(1, 23) = 10.291, p = 0.004, \eta_p^2 = 0.309$), and novelty ($F(1, 23) = 10.505, p = 0.004, \eta_p^2 = 0.314$). No significant interaction effects between agent embodiment and thematic closeness were found. The results are illustrated in Figure 2d.

After a participant experienced all conditions, he was asked for a subjective ranking of the four different guides. Embodied guides were preferred by most of the participants, with 6 votes for the generic museum guide and 14 votes for the astronaut. In comparison, the unembodied generic guide took last place for 12 and the unembodied astronaut for 9 of the participants.

In our pre-study, a participant pointed out an unfair inequality between guides, because he perceived the condition with an embodied astronaut to be the only one dubbed by a real person, while the others were assumed to be generated by a text-to-speech engine. Since even the astronaut guides with and without body were

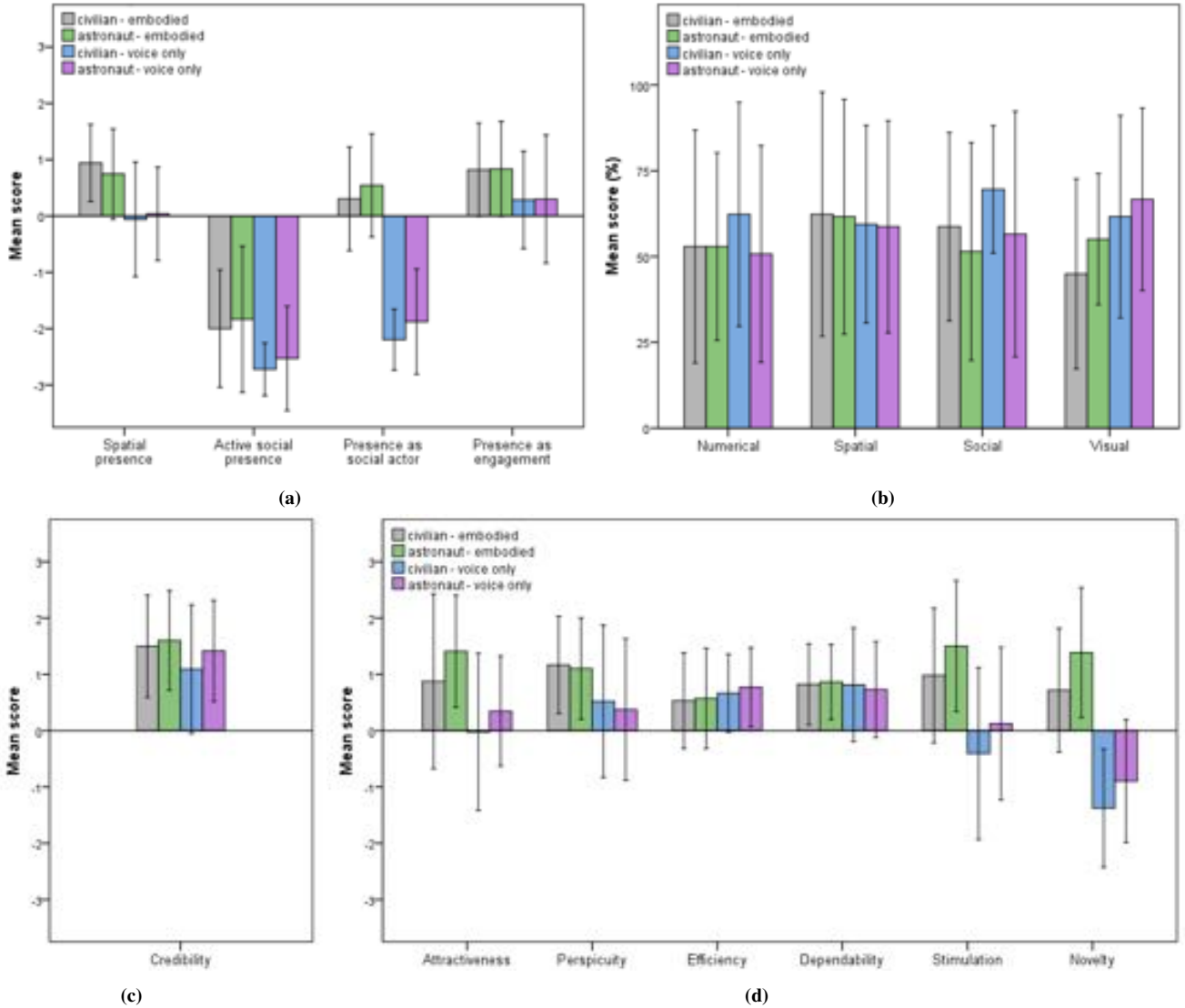


Figure 2: Pooled results of (a) different presence measures, (b) learning results in four categories, (c) agent credibility, and (d) six dimensions of user experience. The vertical bars show the standard deviation.

rated differently, although the same artificially generated voice was used for both of them, we decided to pursue investigations on this aspect in the main study. For each guide participants had to decide whether the spoken text seemed to be produced by a text-to-speech engine or by a real speaker. For the unembodied astronaut, 45.8% of the participants assumed that the agent was synchronized by a real person. For the embodied astronaut, this was the case for even 62.5% of all participants. In contrast, the option of a real speaker was chosen by 37.5% of the participants for the embodied generic guide, and only by 33.3% for the unembodied generic guide.

3.5. Discussion

Even though our exemplary museum application did not include any forms of active interaction between the participants and the

guide, the agent's embodiment had a positive effect on all measured presence dimensions. Through the presence of a second individual within the CAVE, participants felt significantly more spatially involved in the virtual environment. Participants also reported that the embodied guides caused more emotional responses such as laughing or smiling. In general, there was only little active interpersonal communication between users and guides in all conditions, however, this could be regulated by the introduction of additional interaction mechanisms such as voice commands. Whether this is desirable strongly depends on the application itself. In public settings such as a museum, speaking with a virtual agent may make users feel uncomfortable. In contrast, speaking with a personal assistant at home is already the custom and generally accepted. The most remarkable difference between embodied and unembodied guides can be observed in the scores of presence as social actor, sometimes

also referred to as parasocial interaction. This measure of presence contains items that are related to crossing the border between the actual physical environment and the mediated environment in order to interact with the agent in real time [LDC*00]. Higher scores for embodied guides indicate that participants felt that their presence was noted by the agent and that he was establishing a connection to them. Although no complex reactions of the agent to the user's behavior were implemented, a feature as simple as making eye contact seems to be an effective method to create a sense of responsiveness and intimacy. Not only the agent's embodiment but also his thematic closeness had a main effect on presence as social factor. This positive effect could be caused by the first-person perspective of the astronaut, since the guide was not only imparting knowledge but was inviting the user to take part in his personal story.

Regarding credibility, all guides got mean scores in the upper range of the 7-point scale. Besides the realism of guides this could also be attributed to the fact that users do not expect museum guides to lie to them about the chronological order of historical events. Nevertheless, we found a significant effect of agent embodiment on the perceived credibility, indicating that embodied guides seemed to be even more competent and trustworthy.

Despite the exam was announced beforehand to the participants of the study, we expected different learning results for the four types of agents, in particular with regard to the different categories of information. However, this hypothesis could be confirmed only to some extent. Visual details such as the color of specific objects could be remembered better in conditions with a voice-only guide than in scenes with an embodied guide. We expected this outcome, since users tend to follow the agent's lip movements in the embodied condition and therefore could be more distracted from the actual scene. On the other hand, we hypothesized a positive effect of embodiment on the memorization of spatial information, however, such an effect could not be found in the data. In contrast to the results of the objective oral exam, participants subjectively perceived their gain of knowledge to be higher in the conditions with embodied guides, in particular in the long term. Indeed, the involvement of multiple modalities in the learning process as well as an increased presence in virtual environments were related to better learning results in previous studies [Mik06]. A follow-up study that focuses on long-term effects of learning could resolve the question whether the subjective impression of participants can be sustained by an objective test.

We also expected the scores for social questions to be higher for guides with a personal connection to the stories. While no significant effect was found between the generic museum guide and the astronaut, Figure 2b even indicates a trend in favor of the generic guide. The comment section of the questionnaires could give some indication of possible reasons for the observed behavior. Some participants stated that the astronaut was harder to understand due to the applied radio transmission effect.

Besides the problems in understanding the astronaut due to the added distortions, it was also mentioned that this effect made the astronaut sound more realistic than the generic guide. This impression was also confirmed by the responses to the question whether the audio was generated with a text-to speech engine or spoken by a real person. Besides the thematic closeness of the agent, his em-

bodiment also affected the perceived realism of his voice in a positive manner. In spite of identical audio tracks, the presence of an embodied agent seems to distract the user from artifacts of speech synthesis and made the voice sound more natural. Therefore, the embodied astronaut was perceived to have a real voice by the majority of participants.

The perceived realism of the astronaut could also contribute to his positive reception by the participants of the study. In the usability questionnaires, the astronaut guides were rated as significantly more attractive, exciting and motivating, as well as innovative and creative. This is also true for embodied guides in comparison to guides with voice only. These results indicate that the extra effort that has to be made in order to implement a customized agent could be worthwhile to increase user satisfaction and improve the overall user experience.

4. Conclusion

In this paper, we presented a user study that provides some insight on the effects of embodiment and thematic closeness of virtual agents on practice-oriented aspects such as social presence, agent credibility, learning results, and user experience. In particular, we wanted to investigate whether the costly and time-consuming implementation of embodied agents and their customization to a specific application give a competitive edge over common AI agents with audio only.

With a museum scenario we chose a domain, in which audio guides are common for providing people with information. We compared these well-known audio guides with embodied guides and found significant differences in spatial and social presence, credibility, as well as perceived attractiveness, perspicuity, stimulation, and novelty. These positive effects could justify the extra effort for an embodied agent, especially in public installations, where the user experience is of top priority and technical requirements such as an additional projector can be met more easily.

Regarding the thematic closeness of the guide, we found an improvement of user experience in the dimensions attractiveness, stimulation, and novelty. We also hypothesized, that historical guides with a personal connection to the told story are more credible and boost the learning results, because users are able to emphasize with their feelings and emotionally engage with them. While credibility was slightly higher for the astronaut guide than for a generic museum guide, we could not find positive effects of thematic closeness on the learning results in our user study. However, some participants reported that their motivation to explore the VR environment predominated their interest in the Apollo 11 mission. On the one hand, this should be different for users who visit a museum by their own choice, because they are interested in the museum's theme. On the other hand, even users who have no thematic interest could be still attracted by the novel technology and therefore there is a bigger chance to learn something new as if they would not pay attention at all.

While one important aspect in the presented study was knowledge transfer, other domains may have different challenges that have to be met by virtual agents. For example, in health care for children a virtual expert such as a doctor could be compared to a

less intimidating agent such as a mascot. In this example, credibility and social presence are still of great importance, but in addition other variables are emerging, such as the release of fears. Additional studies are necessary to fully answer the questions which agents perform best in different scenarios.

References

- [Ali05] ALIBALI M. W.: Gesture in Spatial Cognition: Expressing, Communicating, and Thinking About Spatial Information. *Spatial cognition and computation* 5, 4 (2005), 307–331. 1
- [Bla02] BLASCOVICH J.: Social Influence within Immersive Virtual Environments. In *The Social Life of Avatars*, Schroeder R., (Ed.), Computer Supported Cooperative Work. Springer London, 2002, pp. 127–145. 2
- [BLB*02] BLASCOVICH J., LOOMIS J., BEALL A. C., SWINTH K. R., HOYT C. L., BAILENSON J. N.: Immersive Virtual Environment Technology as a Methodological Tool for Social Psychology. *Psychological Inquiry* 13, 2 (apr 2002), 103–124. 2
- [CRW*13] CHUAH J. H., ROBB A., WHITE C., WENDLING A., LAMPOTANG S., KOPPER R., LOK B.: Exploring Agent Physicality and Social Presence for Medical Team Training. *Presence* 22, 2 (2013), 141–170. 2
- [DNP11] DEMEURE V., NIEWIADOMSKI R., PELACHAUD C.: How Is Believability of a Virtual Agent Related to Warmth, Competence, Personification, and Embodiment? *Presence: Teleop Virt* 20, 5 (2011), 431–448. 2
- [Dv00] DEHN D., VAN MULKEN S.: The Impact of Animated Interface Agents: A Review of Empirical Research. *International Journal of Human-Computer Studies* 52 (2000), 1–22. 2
- [FAJ*14] FOX J., AHN S. J., JANSSEN J. H., YEYKELIS L., SEGOVIA K. Y., BAILENSON J. N.: Avatars Versus Agents: A Meta-Analysis Quantifying the Effect of Agency on Social Influence. *Human-Computer Interaction* 30, 5 (2014), 401–432. 2
- [GPS72] GLASS G. V., PECKHAM P. D., SANDERS J. R.: Consequences of Failure to Meet Assumptions Underlying the Fixed Effects Analyses of Variance and Covariance. *Review of educational research* 42, 3 (1972), 237–288. 5
- [HB04] HARMS C., BIOCCA F.: Internal Consistency and Reliability of the Networked Minds Measure of Social Presence. In *Annual International Presence Workshop* (2004), pp. 246–251. 2
- [HCO*11] HOLZ T., CAMPBELL A. G., O'HARE G. M. P., STAFFORD J. W., MARTIN A., DRAGONE M.: MiRA - Mixed Reality Agents. *International Journal of Human-Computer Studies* 69, 4 (2011), 251–268. 2
- [HDO09] HOLZ T., DRAGONE M., O'HARE G. M. P.: Where Robots and Virtual Agents Meet: A Survey of Social Interaction Research Across Milgram's Reality-Virtuality Continuum. *International Journal of Social Robotics* 1, 1 (2009), 83–93. 2
- [HRHO92] HARWELL M. R., RUBINSTEIN E. N., HAYES W. S., OLDS C. C.: Summarizing Monte Carlo Results in Methodological Research: The One- and Two-Factor Fixed Effects ANOVA Cases. *Journal of educational statistics* 17, 4 (1992), 315–339. 5
- [KBB*18] KIM K., BILLINGHURST M., BRUDER G., DUH H. B.-L., WELCH G.: Revisiting Trends in Augmented Reality Research: A Review of the 2nd Decade of ISMAR (2008-2017). *IEEE TVCG* (2018), 1–16. 1, 2
- [KBMW16] KIM K., BRUDER G., MALONEY D., WELCH G.: The Influence of Real Human Personality on Social Presence with a Virtual Human in Augmented Reality. In *Proc. of ICAT-EGVE* (2016), pp. 115–122. 3
- [KBW17] KIM K., BRUDER G., WELCH G.: Exploring the Effects of Observed Physicality Conflicts on Real-Virtual Human Interaction in Augmented Reality. In *ACM VRST* (2017), pp. 1–7. 2
- [KMB*17] KIM K., MALONEY D., BRUDER G., BAILENSON J. N., WELCH G. F.: The Effects of Virtual Human's Spatial and Behavioral Coherence with Physical Objects on Social Presence in AR. *Computer Animation and Virtual Worlds* 28, 3–4 (2017), e1771. 2
- [KNB*17] KIM K., NAGENDRAN A., BAILENSON J. N., RAJ A., BRUDER G., LEE M., SCHUBERT R., YAN X., WELCH G. F.: A Large-Scale Study of Surrogate Physicality and Gesturing on Human-Surrogate Interactions in a Public Space. *Front. Robot. AI* 4, 32 (2017), 1–20. 2
- [LBHW18] LEE M., BRUDER G., HÖLLERER T., WELCH G.: Effects of Unaugmented Periphery and Vibrotactile Feedback on Proxemics with Virtual Humans in AR. *IEEE TVCG* 24, 4 (2018), 1525–1534. 2
- [LDC*00] LOMBARD M., DITTON T. B., CRANE D., DAVIS B., GIL-EGUI G., HORVATH K., ROSSMAN J., PARK S.: Measuring Presence: A Literature-based Approach to the Development of a Standardized Paper-and-Pencil Instrument. In *PRESENCE 2000* (2000), vol. 240, pp. 2–4. 7
- [LDW09] LOMBARD M., DITTON T. B., WEINSTEIN L.: Measuring Presence: The Temple Presence Inventory. In *Proc. of PRESENCE 2009* (2009), pp. 1–15. 4, 5
- [LHS08] LAUGWITZ B., HELD T., SCHREPP M.: Construction and Evaluation of a User Experience Questionnaire. In *USAB* (2008), Springer, pp. 63–76. 5
- [LKK96] LIX L. M., KESELMAN J. C., KESELMAN H.: Consequences of Assumption Violations Revisited: A Quantitative Review of Alternatives to the One-Way Analysis of Variance F Test. *Review of educational research* 66, 4 (1996), 579–619. 5
- [Mik06] MIKROPOULOS T. A.: Presence: a Unique Characteristic in Educational Virtual Environments. *Virtual Reality* 10, 3–4 (2006), 197–206. 7
- [MNW14] MCGLOIN R., NOWAK K. L., WATT J.: Avatars and Expectations: Influencing Perceptions of Trustworthiness in an Online Consumer Setting. *Psychology Journal* 12 (2014). 5
- [MTPC08] MAGNENAT-THALMANN N., PAPAGIANNAKIS G., CHAUDHURI P.: Applications of Interactive Virtual Humans in Mobile Augmented Reality. In *Encyclopedia of Multimedia*, Furht B., (Ed.), 2nd ed. Springer, 2008, pp. 362–368. 2
- [NB03] NOWAK K., BIOCCA F.: The Effect of the Agency and Anthropomorphism on Users' Sense of Telepresence, Copresence, and Social Presence in Virtual Environments. *Presence* 12, 5 (2003), 481–494. 2
- [NKH*18] NOROUZI N., KIM K., HOCHREITER J., LEE M., DAHER S., BRUDER G., WELCH G.: A Systematic Survey of 15 Years of User Studies Published in the Intelligent Virtual Agents Conference. In *Proc. of IVA* (2018). 1, 2
- [ODK*12] OBAID M., DAMIAN I., KISTLER F., ENDRASS B., WAGNER J., ANDRÉ E.: Cultural Behaviors of Virtual Agents in an Augmented Reality Environment. In *International Conference on Intelligent Virtual Agents* (2012), Nakano Y., Neff M., Paiva A., Walker M., (Eds.), vol. 7502 of *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, pp. 412–418. 2
- [ONP11] OBAID M., NIEWIADOMSKI R., PELACHAUD C.: Perception of Spatial Relations and of Coexistence with Virtual Agents. In *International Conference on Intelligent Virtual Agents*, Vilhjálmsson H., Kopp S., Marsella S., Thórisson K., (Eds.), vol. 6895 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2011, pp. 363–369. 2
- [Sla09] SLATER M.: Place Illusion and Plausibility can Lead to Realistic Behaviour in Immersive Virtual Environments. *Philos Trans R Soc Lond B Biol Sci* 364, 1535 (2009), 3549–3557. 2
- [WLK*15] WANG Y., LUCAS G., KHOOSHABEH P., DE MELO C. M., GRATCH J.: Effects of Emotional Expressions on Persuasion. *Social Influence* 10, 4 (2015), 236–249. 2
- [YBR07] YEE N., BAILENSON J. N., RICKERTSEN K.: A Meta-Analysis of the Impact of the Inclusion and Realism of Human-Like Faces on User Experiences in Interfaces. In *Proc. of ACM CHI* (2007), pp. 1–10. 2