MagicARTs: On the Design of Social VR Experiences

Maria Vayanou[1], Kyriakos Christodoulou[1], Akrivi Katifori[2] and Yannis Ioannidis[2]

[1]Department of Informatics and Telecommunications, National & Kapodistrian University of Athens, Athens, Greece
[2]ATHENA Research and Innovation Center, Athens, Greece

Abstract

In this work we use the MagicARTs social VR experience for a user study with 15 participants in 7 groups, exploring different social features and interactions in a virtual fine-arts gallery context. The analysis of results revealed that these features seemed to effectively contribute to the overall engagement of the participants with the cultural content and one another. Social VR can be a key motivation for user engagement and create opportunities to transform the VR gallery into a social space that facilitates joint reflection and conversation, inspired by art.

Keywords

Virtual reality, Virtual gallery, Social VR, Social interaction

1. Introduction

The use of Virtual Reality (VR) to support cultural visits, either within the physical environment of cultural institutions [1] or remotely [**?**], has been extensively studied, from the perspective of user engagement and learning [2] to embodied interaction [3]. VR for cultural heritage has been explored in combination with other technologies in the reality-virtuality continuum [2, 4], often highlighting the value of social interactions or investigating in what ways and conditions the users initiate conversations and engage in social interactions [4].

This work addresses remote cultural visits that take place in Virtual 3D Environments (VEs), using exclusively immersive VR equipment. It focuses on the social dimension of the visit, an aspect that is highlighted as one of the most important experience design pillars. However, it is still notably underexplored in the virtual museums world, as VR technology has been applied mainly to provide single user experiences [5].

First, we present MagicARTs, a multi-user VE that features a collection of Vincent Van Gogh's paintings in a virtual gallery hall, along with informative and artistic multimedia content. MagicARTs does not implement a persistent virtual world (as Second Life and its immersive VR descendant, Sansar). Following game-based approaches, it provides private "rooms" that the group members may join together, in absence of strangers. There are no specific joint activities offered in these rooms; no puzzle to solve, no game to play, no storyline to imply a particular exploration path within the virtual gallery. Even though related literature has shown that game-based activities or concrete tasks in cultural VEs lead to increased user engagement

AVI-CH 2022 Workshop on Advanced Visual Interfaces and Interactions in Cultural Heritage. June 06, 2022. Rome, Italy
vayanou@di.uoa.gr (M. Vayanou); sdi1300211@di.uoa.gr (K. Christodoulou); vivi@di.uoa.gr (A. Katifori); yannis@di.uoa.gr (Y. Ioannidis)

[6, 3], we believe that the "free visit" metaphor has much to offer, and sociality may be one of the ways to make it more enjoyable.

Moving in that direction, we describe the social design elements of MagicARTs, presenting in detail the output of each user action in the VE and discussing how it is delivered to different users - using private or social distribution mechanisms. Then we present the results of a user study with 15 participants, who visited MagicARTS in small groups. We report the effects of social interactions in overall participants' experience, while also examining whether the provided social features were noticed, how they were used and overall perceived. Such issues are commonly overlooked [7], yet they provide valuable insights for shaping and fine-tuning the experience design pf VR museums.

2. Related Work

The pioneering work in "Virtual Leonardo" [8], a desktop-based multi-user virtual museum, showed that the social dimension of the virtual visiting experience was particularly appreciated. However, two main issues (related to this work) were observed: a) the visitors communicated by writing text messages, which often resulted in "hanging avatars" within the VE, and b) the visitors did not really know how and why to interact - a rather critical issue that is often addressed by providing games or specific activities in the VE, hence more or less guiding social interaction through experience design [6].

Over 20 years have passed since the development of "Virtual Leonardo" and the technological advances have shaped a new landscape for VR applications. Nowadays, voice-communication is easily enabled and hand-gestures, whose role and value has been recognized as "not merely add-ons to language", but rather "a fundamental part of it" [9] are effectively captured with common hand controllers, and effectively reflected in the VE. However, the design of social VR applications [10] still raises significant challenges. Apart from gestures, the indirect, non-verbal communication channel between two humans is rather poorly reflected in the VE: face expressions, eye-gaze, attention focus and body posture are important elements of social interaction that are missing. The research work in [11] attempts to signal such social cues in a virtual gallery environment using visual feedback augmentations - for instance, "eye contact" between two virtual selves is visualized with floating bubbles. In a similar direction, MagicARTS proposes an interaction design of enhanced social cues to promote social encounters in the gallery.

3. MagicARTS Design

The MagicARTS gallery is the descendant of a previous desktop-based VR work [12] and it is based on a simple space design of a large rectangular hall. Upon entering the gallery, the artworks are exhibited on the wall. The gallery follows a simple low poly design, with no ceiling and a virtual sun acting as the natural light source. The artworks thus become a natural focus of the room, due to their photorealistic presentation and vivid colors. There are no additional cultural resources visible upon entering. These are gradually activated by the visitors, as they

Table 1
Summary of User Interactions with Elements in MagicARTS

	Elements of Interaction	Action	Output	Output Distribution	Social Activity Cues
1	Painting image	Select in VE	Change the VE state: Activate elements #2-4	Social - Symmetrical	Exploration sound effect
2	Text Narrative	Select in VE	Start/pause audio narration	Private	Border highlighting of related painting & headset on avatar
3	Video Screen	Select in VE	Start/pause video presentation	Private	Border highlighting of related video screen & headset & HMD on avatar
4	Emoticons	Select in VE	Short-term visual effect	Social – Asymmetrical (performer vs spectators)	Emoticon declaration sound effect
5	Animated painting presentation	Approach artwork (in red semi-circle)	Start video projection over the painting	Social - Symmetrical (sound-speaker metaphor)	
6	All users	talk	voice-communication	Social - Symmetrical (voice-call metaphor)	
7	All users	see	avatar representation	Social - Symmetrical	

interact with the virtual gallery. A number of actions are available to the participants, each with a different effect on the VE (Table 1).

In the remainder of this section we present the VE actions and their output and then focus on the social features.

3.1. Visitor Actions and their Output in the VE

3.1.1. Selecting Virtual Objects

At the beginning, the main elements of interaction are the digital representations of the artworks. When a visitor selects a painting to further explore it, additional elements get displayed on the wall around it, simulating real-life exhibition practices: a related narrative text shows up on its right (along with a headset visual indication), a set of 5 emoticons on the top, and in some cases a black "video screen" appears on its left, featuring documentary films, interviews by museum experts or informative presentations relating to the painting. These complementary elements become visible to all visitors in the VE and from this point and on, they remain permanently on the gallery's wall (until the end of the group visit session). In this way, the virtual gallery environment provides a "joint exploration trail": if the wall around a painting is empty, one can infer that none of the group members has explored that painting. The visitors may then interact with each of these elements as follows: i) select the text narrative to activate its audio narration, and then select it again to pause/resume it (there is no other media handling interface provided), ii) select the video screen to activate a related informative video presentation and then, as with audio, reselect to pause/resume it, iii) select one out of the 5 emoticons on top of the painting to denote the feelings they may have about it.

While the first two actions have a clear objective and closely match the visiting activities that may typically take place during an on-site cultural visit, the third one is inspired from the digital world. Users have nowadays been accustomed to denoting not only their likes and dislikes, but also a wider range of emotional responses over a variety of stimuli - over social networking platforms or even instant text messages – "marking" the related elements of interaction with different types of emoticons. But how should emotional responses to paintings be reflected in

Figure 1: Visual effects following emoticon declaration from each side: spectator (left) and performer (right)

the 3D world? Should we follow a traditional 2D visualization scheme that the users are already familiar with, or maybe move towards a new 3D-oriented approach?

The rationale for introducing this element in the design of MagicARTs was that visitors' declaration of emotional responses would potentially stimulate perspective sharing and discussions over the paintings, which is one of our main design objectives. In that direction, and conceiving it as an action of revealing personal reflections, we designed the output of emoticon selection as "thought bubbles" that are displayed in the VE over the avatars' head (and vanish after a while) - see Figure 1.

3.1.2. Approaching Virtual Objects

Apart from the informative videos, there are 4 paintings in the gallery for which a fictional animated version of the artwork is provided through video presentations, featuring the movement of depicted figures or scenery (e.g. persons walking, lights blinking, waves moving, etc.), adding a sense of depth, or even showing how the depicted spaces could look from their inside (if they were 3D). These videos include mainly background music and no voice narration; instead of delivering information they aspire to 'bring life' to the paintings. As such, in the MagicARTs gallery such videos get projected over the artworks, taking their place on the wall, and are activated automatically when a visitor approaches the relevant painting. The activation perimeter is denoted over the gallery's floor with a red semi-circle surrounding the corresponding paintings, signaling this indirect type of location-based interaction in the VE. When there is no visitor within the activation perimeter, the video stops, and the original image of the painting takes its place again on the wall. In essence, these areas are conceived as "Shared Projection Areas" and are inspired from real world cultural exhibitions with digital installations. Joint attention areas are often used to promote social interaction., so we decided to explore this metaphor.

3.2. Social Features

3.2.1. User representation

Visitors are represented in the virtual gallery environment with identical, simplistic shape-based avatars having a head, torso and hands (see Figure 2), along with a 2D name tag displayed on

Figure 2: User representation in MagicARTS gallery

top of each avatar's head.

The movement of head and hands is tracked with the VIVE equipment and effectively reflected in the VE. On the other hand, there is no actual tracking of the user's torso, so its virtual counterpart is attached at the central axis of the virtual head, thus practically following the users' head movement.

User representation is a broader topic in VR research, in many ways affecting socialization along with other important aspects of user experience, including presence and immersion. In this work, we decided to leave this aspect out of investigation, following the example of [11]. We focus on visitors' interactions with the gallery content and explore how these are socially mediated within the cultural VE. The investigation of more elaborate user representation approaches (from customized colorings to cartoon-like or even more realistic avatars) is a pertinent topic for future work (see limitations section).

3.2.2. Voice communication between users

Voice communication between the group members is supported in the VE at all times. Regarding the handling of its volume, we examined two different approaches: either mimic real-life world conditions and have the volume decreasing as the distance between the virtual selves increases, or follow an "in-voice-call" metaphor and keep the volume of the voice communication channel stable. Aiming to promote social interactions between small groups of visitors (pairs and triads), we opted for the latter.

3.2.3. Social Output Distribution

The visitors share the same world (i.e. their avatars co-exist in the same gallery hall in which one can see each other in real time) and make individual actions within this world. The output of each action may be either *private*, i.e. experienced only by the user who performs the action, or *social*, in the sense that somehow affects not only the performer, but has some output on other users as well.

Output distribution may be "the same" for everybody, in which case we call it *symmetrical*, but it may also vary, being *asymmetrical* at different levels. For instance, the action of selecting an unexplored painting shapes the attributes of the shared VE itself, hence its output distribution is symmetrical. The same holds for the action of activating the "Alive" video presentations of a painting (when entering its red semi-circle area): the video is projected over the wall and all users may see it, even if they are not within the activation perimeter.

It is worth noting that, by adopting a "sound speaker metaphor", the volume of the audio depends on the distance of each avatar from the video-projection space. As a result, users who happen to be close to it will hear its background music in higher volume than users who are far away. So, one could say that there is some level of asymmetry here, yet in this case the multimedia content distributed between all users remains the same. Furthermore, some level of asymmetry is inherently underlying a multi-user experience of shared 3D worlds (from personal avatar perspectives), hence we assume that distance-based variances in volume levels do not constitute asymmetry.

On the other hand, the visual effect following the action of emoticon selection differentiates between spectators and performers. Spectators see a "thought bubble" over the avatar of the user who made the action, but due to the placement of the bubble over the avatar's head, the performer was not expected to see it (unless they turned their head up – which is not intuitive). So a different visualization is distributed in the performers' side instead: an animated "raining" effect appears in their viewpoint, having several instance of the selected emoticon crossing it vertically, as if "falling from the sky" (see a snapshot in Figure 1).

3.2.4. Private Output Distribution

Following common museum practices, the MagicARTS gallery offers informative content for the paintings, through text narrations and video presentations. One of the main reflection points in our design was about whether this content should be distributed privately, and thus consumed by each user separately, or instead be socially (and symmetrically) distributed - following again the sound-speaker metaphor. Since our design accounts for small groups (i.e. pairs or triads), we believe that both are promising alternatives.

Private media distribution would allow the group members to diverge from each other and adopt different paces if they wish to, enabling them to make personal selections and focus on different aspects and artworks. For instance, when a visitor activates the audio narration for a painting, the related audio is delivered exclusively to the user who selected it - not to everybody in the VE. While doing so, other visitors may start interacting with the same object: at their side, the audio narration starts from scratch and a different "progress state" is maintained per user. In this way, each visitor can pause the narration privately at any time, and resume it form the

point they had (personally) left it. However, this empowerment may often come at the expense of social interaction, favoring the individual experience over the social dimension of a group visit.

On the other hand, social distribution of media content would require visitors to significantly coordinate their actions and more or less visit the gallery together, while potentially leading to uncomfortable situations or posing issues of control. Consider for instance the following scenario: one user starts the audio narration for a painting and another user starts its video presentation right next to it. If no special handling is supported by the VE software, then both audios would play concurrently, which would be rather disorienting and annoying for both users. Several levels of handling mechanisms could be implemented to that, from simply stopping one video before the other starts to using sophisticated locking techniques – yet these come with other issues (which are out of the scope of this paper). Here we describe our approach to support private media distribution in the design of MagicARTs by enriching it with activity cues that signal private content consumption to other users.

3.2.5. Social Activity Cues

User actions in a VE are commonly accompanied by audio or/and visual cues, indicating the object of interaction, or simply signaling the completion of actions. Such cues are particularly helpful when private distribution is employed in multi-user environments, as they promote awareness of other users' actions and state.

In this direction, while the visitors of MagicARTs are listening to an audio narration or watching a video presentation, the border of the related virtual object is highlighted (in green), indicating the cultural resource they are currently engaged with. The visual cue is visible to all users, thus hinting when moments of private distribution are encountered. But is this enough?

In real world group visiting conditions, one may quite easily perceive the attention focus of their companions, observe their eye-gaze, face expressions, body posture and understand whether they are actively engaged with some activity or not. Some visitors use this information to coordinate their social interactions – e.g. avoiding to interrupt such moments of engagement and wait until their companions disengage from an activity, in order to talk to them. Unfortunately, this indirect social communication channel is not available in a VE. Furthermore, the virtual selves may be located far away from the object of interaction (e.g. one may be freely moving around while listening an audio narration), so the visual highlighting of related objects may often go unnoticed.

To cope with the disparate locations of avatars and their interaction objects, some VEs also employ a virtual tray that visualizes the connection between the two. Aiming to stress out when such moments of personal engagement occur (rather than the identity of the related content), MagicARTs explores a different approach. To signal users' engagement with narrative audio-visual presentations, a real-life metaphor is transferred in the virtual world: when a visitor is listening to an audio narration, virtual headphones get displayed on their avatar's head, while when watching a video presentation a virtual HMD is employed as well (Figure 3).

Apart from using virtual visual elements to socially signaling private output distribution phases, MagicARTs also implements sound-effect based cues for actions whose output is socially distributed. In particular, artwork and emoticon selection actions have visual outputs only

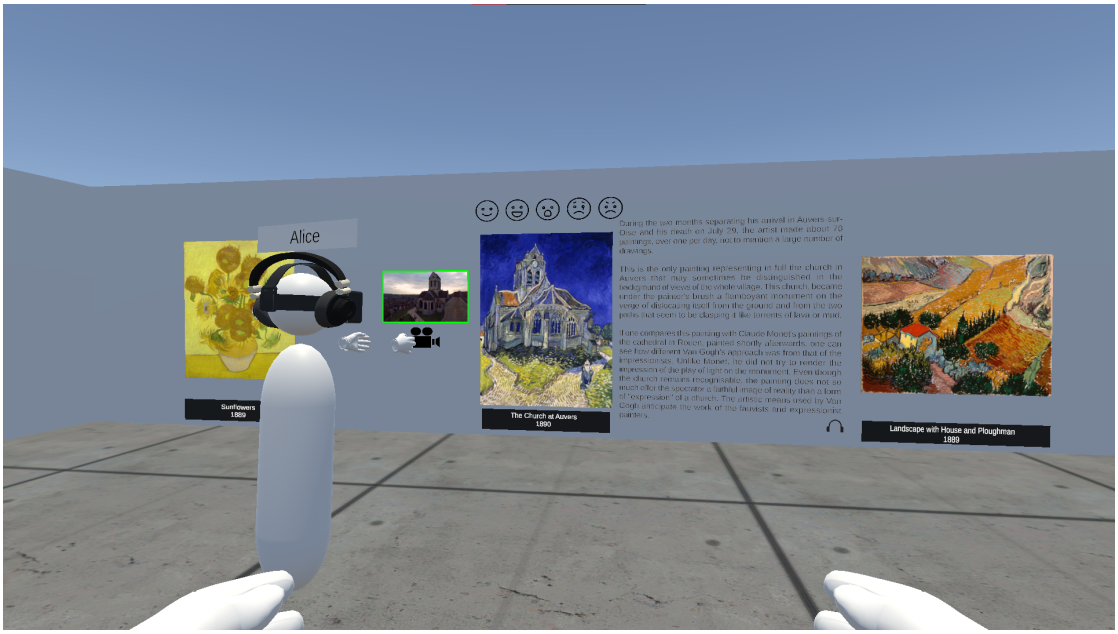

Figure 3: HMD and headset display on avatar while watching a video presentation - highlighted in green color

(long-term and short term correspondingly). In a 3D world, users may be frequently looking in a different direction than the performer's avatar, so audio cues are used along with the visual output to indicate that some actions are currently taking place by other users in the VE, somehow altering its state.

4. User Study

Focusing on the social dimension of cultural virtual visits, we performed a user study with 15 adult participants in the immersive VR environment of MagicARTs, to examine whether and how the social features of our design were perceived, and used in practice, by the participants. First, we assess matters related to overall user enjoyment and social presence within the VE, and then go into more depth for each one of the social features offered in the VE, setting a series of explorative questions: Did the participants notice the actions' social distribution mechanisms? Did they notice the social activity cues? Did these somehow affect their social behavior and how? Did the participants face any problems relating to the social features of the design?

4.1. Participants

An open invitation was sent to students of the first author's university department to participate in an evaluation of an immersive VR experience. They were also invited to bring with them friends or family to participate together in the experience. 15 users divided in 7 groups participated in the study, which included 3 couples, one parent and daughter and 3 groups of

friends. 6 groups were of 2 users and one of 3. The average participant age was 28.4, with the youngest participant being 24 years old and the eldest 45. There were 9 men and 6 women.

11 participants had no prior experience with desktop-based VR applications while the rest had a littler (through computer games). 8 participants reported some prior experience with immersive VR, for entertainment and gaming. In terms of their interest in modern art, in a scale of 1 (not all) to 7 (very much), the responses are almost equally distributed between 2 and 6.

4.2. Process

The study was organized to take place at a VR center. The members of a group were invited to arrive at a designated date and time together. Some days before the evaluation they were asked to fill in a user profile questionnaire.

When they arrived at the VR Center, they were welcomed and led to a waiting room where they were offered refreshments and were informed of the objectives of the study and given the opportunity to read the relevant information sheet. They were then asked to sign the consent form. Next they answered a set of questions about their current physical state, as a baseline to compare with potential motion sickness symptoms afterwards.

Then the members of the group were led to separate rooms where the VR equipment was available. The HTC Vive VR equipment has been used, including the HMD and controllers. The participants were supported to put on the equipment and once they were ready, the experience was initiated. The evaluators explained that the participants could inform them at any time if they were uncomfortable or they wished to end the experience. The participants could stay in the VE for as long as they wished and did not need to stop the experience together - if one group member wished to leave the VE but the other wanted to continue, they were free to do so. After they were given a set of simple instructions, they were ready to start the experience.

Directly after the experience they were asked a motion sickness related questionnaire and then they were led to the waiting room and offered refreshments. After resting for a few minutes, they were asked to fill in a questionnaire and then interviewed by the evaluator. The questionnaire is presented in more detail in the following section.

4.3. Study method and instruments

For the study reported in this work, a mixed methods approach has been applied, combining qualitative and quantitative measures. More specifically, we employed: (1) a questionnaire filled in individually by the participants in Google Forms format after the VR experience, (2) a semi-structured interview, once the participants filled-in the questionnaire and (3) video recordings of the VR experience from the perspective of all the participants.

The questionnaire consists of 18 items: 16 in 7-point Likert scale and two open questions. It is organized in two parts, “Overall assessment” and “Sociality assessment”. The first part examines aspects of enjoyment and willingness to revisit the virtual gallery (with the same or different artwork collections). It includes 3 items, along with two open questions to investigate the main strengths and weaknesses of the experience (*Describe 3 aspects you liked the most/ did not like*). The “Social assessment” part contains 13 elements, of which 11 are adapted statements from the self-reporting items of the Networked Minds [13] and GEQ [14] (social presence

Table 2
Evaluation: Average and Standard Deviation Scores

N.	Questionnaire items	AV / STD
	Overall assessment	
1	Did you like the experience?	6.33 / 0.98
2	Would you like to invite your friends or family and revisit this gallery with them?	6.00 / 1.07
3	Would you like to visit more galleries like this (with different artworks)?	5.87 / 0.99
	Sociality assessment	
1	I often felt as if I was all alone.	3.00 / 1.89
2	I tended to ignore my companion.	2.73 / 1.83
3	I paid close attention to my companion.	5.20 / 1.57
4	I was often aware of my companion's location in the Virtual Gallery Environment.	5.53 / 1.46
5	I was often aware of what my companion was doing in the VE (e.g. which painting/video they were watching).	5.00 / 2.07
6	What my companion did affected what I did.	4.67 / 1.72
7	I often tried to approach my companion's avatar in the virtual environment.	5.67 / 1.40
8	I often tried to remain close to my companion's avatar in the VE.	5.53 / 1.30
9	I felt connected to my companion.	5.27 / 1.39
10	I found it enjoyable to be in the Virtual Gallery Environment with my companion.	6.27 / 0.88
11	I was influenced by my companion's feelings and moods.	5.33 / 1.72
12	Me and my companion often shared thoughts or opinions about the artworks.	5.20 / 1.42
13	Me and my companion often shared feelings evoked by the cultural content.	4.87 / 1.77

module) questionnaires, capturing aspects related to social awareness and attention allocation, behavioral interdependence and psychological involvement. It also includes 2 case specific statements to investigate the cognitive and emotive sharing activities regarding the cultural content that was offered in the virtual gallery (items 12, 13 in Table 2). We added these items because they reflect our two main objectives when designing digitally mediated cultural group visits, either for on-site or remote settings.

For the semi-structured interview, an Interview Guide was used by the evaluators and, upon completion, participants were given the option to fill in an "Open remarks" section.

4.4. Results

The average duration of the VR experience was approximately 26 min., with the shorter visit being around 13 min. and the longest 37 min. In 5 (out 7) groups, when one of the group members asked to terminate the visit their companion(s) asked to do so as well, while in the other two groups the companion continued their visit for a few more minutes (about 4 and 6 min. per case). None of the users asked to finish the experience due to discomfort. After the completion of the experience, 3 of the users reported they felt minor discomfort due to the heaviness of the HMD but they did not feel that it affected their experience.

4.4.1. Overall Evaluation and Sociality Perception

In the "Overall assessment" part of the questionnaire, the experience was overall evaluated positively, and most of the participants expressed interest in visiting the virtual gallery again

(Table 2). In the open questions, 7 participants made references on its social dimension, reporting it as one of the aspects they liked the most (e.g. "*Ability to see your partner*", "*you can communicate with other visitors, different in a good way experience in museums*", "*Companion and dialog*", "*Experienced it with another person*", "*interaction with others*", "*interaction with friends*"), while no negative references were made to that end.

After the semi-structured interview in which the social features were discussed in detail, 4 participants noted down in the "Open remarks" section a few comments or suggestions for improvement, including: the wish to mute others (#2), didn't like that my companion could come so close to me (#1), and wish for joint media consumption with partner (#1).

These issues were discussed in more detail during the interview section. The need for muting was particularly highlighted by one participant of the 3-group session, stating that "*others talking didn't let me concentrate on a painting*". From the video-recordings we observe that in this session, there were plenty of moments in which two of the group members remained in close proximity and attempted to coordinate their visit, while the third member followed a rather private path. On the other hand, one participant explained that they would prefer to hear exactly what their companion was hearing, instead of trying to catch up with them and synchronize their media presentations on their own. In their session's recordings we observe that in this case the pair attempted to explore the gallery's informative content together, discussing what to do at each step, which narration to hear and which presentation to watch, while moving closely together at most times.

Regarding the users' perception on how much they exchanged thoughts or opinions (item 12) and emotions (item 13) over the cultural content with their partner(s), we observe different responses, yet overall indicating a social sign (#10 over neutral in both aspects).

4.4.2. Social Activity Cues

During the semi-structured interview, the participants were asked if they noticed the social activity traces (visual and audio) signaling their companion(s) actions - Table 3 summarizes the number of participants replaying positively / negatively per case. On a first look, these numbers seem to be approximately the same, so what happened? Did 9-10 participants notice them all while the rest missed this aspect entirely? Our results show that this is not the case. There were some "extremes" (#4 noticed all, #1 noticed none) but the rest indicated some combination in between.

In the following, the participants were asked if they used the feedback they got from the visual activity cues (about their companions' state) to somehow coordinate their actions. Most of them (#11) replied negatively (Figure 4).

4.4.3. Shared Projection Areas

With one (strong) exception, the participants appreciated the Shared Projection Areas (Figure 5). When asked about the reasons, two main aspects prevailed: the "Go Alive" sense they created, as well as the element of surprise - something different was happening, something they did not expect, and had to figure out what was going on. The participant who replied negatively commented that he had not understood the purpose of the red semi-circles on the floor. Not

Table 3
Participants' responses over noticing social activity cues

Question	#Pos	#Neg
Did you notice the headphones on your companion's avatar when he/she was listening to an audio piece?	10	5
Did you notice the head mounted display on your companion's avatar when they were watching a video presentation?	9	6
Did you notice the green highlight over the painting that your companion was interacting with?	10	5
Did you notice the sound effects were following some of your companion's actions?	10	5

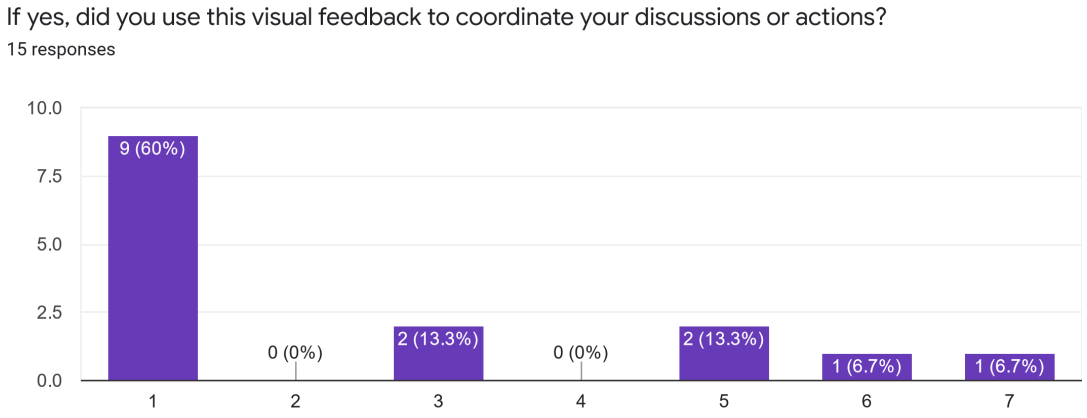

Figure 4: Participants' responses in "Did you use this visual feedback to coordinate your discussions or actions?"

having realized their approach-based activation, it seemed to him as if these media projections were being (de)activated randomly.

Next, the participants were asked if the existence of the Shared Projections Areas affected the coordination with their partner(s), and in what ways. Most of them (#11) replied positively, often recalling cases in which their companion had activated the multimedia projections and they approached them to find out what was happening. Some participants reported that they discussed such videos together - an aspect they enjoyed. One participant pointed out that in one case, the sound of the presentation was too high, and not being able to lower its voice, they had to move out of its perimeter in order to talk about it. Finally, one participant stated that: "*The first time a projection was activated, it triggered my interest so I approached my friend (who was already in the red semi-circle) and actually watched the video together. But that happened only once*". The participant explained that afterwards, he figured out what was happening in the Shared Projection Areas, and did not wish to stop what he was doing at that time in order to re-join his companion.

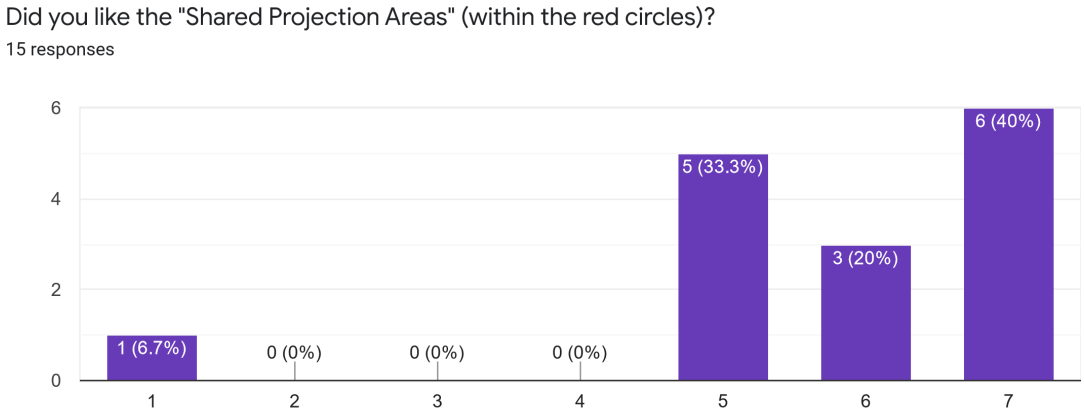

Figure 5: Participants' responses in "Did you like the Shared Projection Areas?"

4.4.4. Emoticons

Most of the participants (#13) actively used the emoticons, performing multiple selection actions, and stated they liked this element. The facilitator asked about their motivation for performing emoticon selection actions. Based on the participants elaborated answers, three main approaches were identified: i) *express appreciation of a painting* (#4), ii) *express emotional response to painting* (#7), iii) *for fun with partner* (#2).

In the first case, the participants explained that they selected only the 'happy' emoticon and never the rest. Perceiving it as a "Like" activity, they performed the action exclusively over the paintings they liked. In the second case, the participants perceived it as a broader expression of their personal feelings about the paintings and selected additional emoticons, often performing the action over every painting. A few participants stated they used the entire emoticons' range, while others used only a subset (e.g. 'happy' and 'sad', or 'happy' and 'surprised').

In the third case, the participants did not express emotional responses about the paintings. They liked the visual effect of their action and used it in a playful way with their companion, either just for fun or to express their feelings about their companion's actions. For example, one participant recalled a case in which she was trying to read a narrative text on the wall but her partner kept moving in front of her, so besides expressing her annoyance verbally, she used the angry emoticon as well.

Regarding the noticeability of emoticons' interaction, one pair of participants reported not to have realized that an action could be performed over the emoticons, and did not interact with this element at all - both stating that they would probably do so if they knew about it. On the other hand, the output of the action was noticed in all groups. The video recordings show that the asymmetric output was identified in all groups sessions (in which the action was performed), and the companions talked about it, e.g. describing to each other the different visual effect they were are seeing at each point and jointly exploring the two different perspectives (as

performers and spectators). Finally, on another aspect, during the interview two participants stated that they would like their emotion declarations to be somehow "registered" and reflected in the VE. One of them explained that, at the beginning of the visit, he thought he was giving a kind of "score" for the paintings. After a while, he realized this was not the case but continued performing the action nevertheless.

5. Discussion

The real-life environment of cultural institutions often poses notable constraints over social interaction during group visits, limiting the extent of conversations between group members. In some cases, fine-art galleries included, there is also a profound "no-talking" visiting culture. On the contrary, in a private VE there are no such limitations. This could be viewed as a positive aspect: If the group members wish, they could talk to each other all the time, without worrying about disturbing other museum visitors or disrespecting the cultural institution's dynamics and rules. The virtual world belongs only to the group. But what happens when the group members have different visiting expectations and styles? Which cultural visit norms are transferred in this new setting, and which are rejected?

This question was a matter of stated disagreement between a pair of participants: Participant A was repeatedly approaching Participant B in the VE, making several attempts to interact, but B was often replying negatively. This went on for a while and at some point Participant B explicitly said: "*Why are you talking to me all the time? One does not talk in art galleries, one watches the artworks*". Participant A shortly paused, but then came back with more questions: "*Why should we not talk?*"... In this case, the social visiting norms of the physical world were maintained in the VE by Participant B, but rejected by Participant A. So, the main question is, how may a virtual gallery design be respectful to both such perspectives?

5.1. Personal moments of reflection vs sociality

The user study results show that the visual activity cues signaling moments of private media consumption were often ignored with regards to social coordination. In addition, two participants expressed their desire to "mute others" for concentrating, as a result of the multiple interruptions experienced in their prior visiting sessions. Safeguarding the personal moments of reflection during a cultural visit is a vital issue, but is muting the best way to go in a private VE?

At the opposite end, one of the participants experienced an entirely different situation, asking why one should "manually" coordinate their actions instead of directly watching and listening to all informative content together. Given that two groups more or less visited the gallery closely together, we wondered whether the interaction metaphors of a VE should not be fixed and same for all, but adaptive, tailoring to different group styles and needs. As a first step towards this direction, we performed a short follow-up interview with the participants, explaining the suggestion made by one of them, and asking whether they would like to visit the virtual gallery in that way. Most participants replied negatively (#10) while only a few were willing to give it a try (#4), however only as long as they had an "escape mechanism" that would allow them to switch back to the current mode if they wished to. The issue of control and autonomy is often

discussed in social, mobile cultural applications and needs more investigation in the case of multi-user virtual museums, where even greater challenges are raised.

5.2. Emotional Sharing

The user study results show that the emoticon selection feature was appreciated and highly used by the participants. We believe that this is due to two main reasons: first, it is an action they are familiar with, practicing it during their everyday social media lives, and second, it moves from spectating to performing, assigning the visitors a more active, contributing role. The suggestion of two participants to have their emotional declarations reflected on the virtual gallery walls is in that direction, and we plan to explore it in our future work.

Our rationale for introducing this element was to trigger emotive sharing and discussions over the paintings. Based on visitor responses in the questionnaire (item 13) this seems to have been accomplished to some extent. The users' perception of the alevel of emotional sharing is an important aspect and we consider it one of the main contributions of this work.

5.3. Shared Projection Areas

Based on the participants' responses, the Shared Projection Areas seem to have effectively served as a social triggering element, creating a joint attention area in the VE and leading to social interactions (including joint video watching and conversations). The use of the sound-speaker metaphor worked well to that end, while also adding an element of surprise: users who were away wondered why they were listening to something they had not activated themselves, and turned over to their companions for answers. We suggest that the "Shared Projection Areas" metaphor has promising potential with regards to cultivating sociality, and is a direction that needs to be further explored in the design of virtual cultural environments.

5.4. Limitations

The MagicARTs social VR gallery was tested with 6 pairs and 1 triad. The later revealed the dangers underlying the use of "voice-call" metaphor to support verbal communication in groups, particularly when moving beyond pairs. The potential hindering of personal moments of reflection and cultural engagement was an issue identified in a few cases of paired use, but greater challenges are posed when attempting to serve groups that have 3 or 4 members. The triad session revealed occasions in which subgroups were formed in the group, while still sharing a single voice-communication channel.

Furthermore, additional visual cues would be required to better reflect the identity of each visitor. For instance, in the current testbed only one color is used for border highlighting, so there is practically no way to tell who is the user that is interacting with a highlighted object. More sophisticated approaches would be required, such as a color coding scheme for user representation (relating avatar color to its activity traces), the addition of "state annotations" over the gallery's walls (e.g. "VisitorName is watching this now") or even activity rays.

6. Conclusions

In this work the MagicARTs immersive social VR is used for a targeted user study with 15 participants in 7 groups, exploring different social interactions features in a virtual gallery context. The analysis of results revealed that these features, especially the emoticons and the shared projection areas, seemed to be effective elements that contribute to the overall engagement of the participants with the cultural content and one another. Social VR can be a key motivation for user engagement with virtual museums, and even attracting new audiences. It also creates opportunities to transform the gallery space into a social space that facilitates joint reflection and conversation, inspired by art.

There are still however many open issues to fully harness the potential of the medium. These include among others, supporting individual user needs and preferences in terms of the desired level of sociality during the virtual visit, ensuring the right balance between sociality and engagement with the artwork as well as understanding the exact function and effect of enhanced social cues in this context.

References

[1] M. Shehade, T. Stylianou-Lambert, Virtual reality in museums: Exploring the experiences of museum professionals, Applied Sciences 10 (2020). URL: https://www.mdpi.com/2076-3417/10/11/4031. doi:10.3390/app10114031.

[2] M. K. Bekele, E. Champion, A comparison of immersive realities and interaction methods: Cultural learning in virtual heritage, Frontiers in Robotics and AI 6 (2019). URL: https://www.frontiersin.org/articles/10.3389/frobt.2019.00091. doi:10.3389/frobt.2019.00091.

[3] B. Flynn, v-embodiment for cultural heritage, in: 2013 Digital Heritage International Congress (DigitalHeritage), volume 1, IEEE, 2013, pp. 347–354. doi:10.1109/DigitalHeritage.2013.6743759.

[4] Y. Li, E. Ch'ng, S. Cai, S. See, Multiuser interaction with hybrid vr and ar for cultural heritage objects, in: 2018 3rd Digital Heritage International Congress (DigitalHERITAGE) held Jointly with 2018 24th International Conference on Virtual Systems & Multimedia (VSMM 2018), IEEE, 2018, pp. 1–8. doi:10.1109/DigitalHeritage.2018.8810126.

[5] M. Vayanou, A. Katifori, A. Chrysanthi, A. Antoniou, Cultural heritage and social experiences in the times of covid 19., in: AVI^{2}CH@ AVI, 2020.

[6] N. Di Blas, P. Paolini, S. Hazan, The see experience: Edutainment in 3d virtual worlds. (2003).

[7] S. Goud, V. Lombardo, Assessment of digital environments for cultural heritage communication, in: AVI2CH 2020 Advanced Visual Interfaces and Interactions in Cultural Heritage, CEUR Workshop Proceedings, 2020, pp. 1–8.

[8] P. Paolini, T. Barbieri, P. Loiudice, F. Alonzo, M. Zanti, G. Gaia, Visiting a museum together: How to share a visit to a virtual world, Journal of the American Society for Information Science 51 (2000) 33–38. doi:10.1002/(SICI)1097-4571(2000)51:1<33::AID-ASI6>3.0.CO;2-I.

[9] S. Kelly, M. Healey, A. Özyürek, J. Holler, The processing of speech, gesture, and action

during language comprehension, Psychonomic bulletin & review 22 (2015) 517–523. doi:10.3758/s13423-014-0681-7.

[10] M. Jonas, S. Said, D. Yu, C. Aiello, N. Furlo, D. Zytko, Towards a taxonomy of social vr application design, in: Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts, CHI PLAY '19 Extended Abstracts, Association for Computing Machinery, New York, NY, USA, 2019, pp. 437–444. URL: https://doi.org/10.1145/3341215.3356271. doi:10.1145/3341215.3356271.

[11] D. Roth, C. Klelnbeck, T. Feigl, C. Mutschler, M. E. Latoschik, Beyond replication: Augmenting social behaviors in multi-user virtual realities, in: 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), IEEE, 2018, pp. 215–222. doi:10.1109/VR.2018.8447550.

[12] K. Christodoulou, M. Vayanou, G. Tsampounaris, Y. Ioannidis, Magicarts: An interactive social journey in the art world, in: Proceedings of the 22nd Pan-Hellenic Conference on Informatics, PCI '18, Association for Computing Machinery, New York, NY, USA, 2018, p. 272–277. URL: https://doi.org/10.1145/3291533.3291579. doi:10.1145/3291533.3291579.

[13] F. Biocca, C. Harms, J. Gregg, The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity, in: 4th annual international workshop on presence, Philadelphia, PA, 2001, pp. 1–9.

[14] W. A. IJsselsteijn, Y. A. De Kort, K. Poels, The game experience questionnaire (2013).