It's Not Always Better When We're Together: Effects of Being Accompanied in Virtual Reality

Rufat Rzayev
University of Regensburg
Regensburg, Germany
rufat.rzayev@ur.de

Florian Habler
University of Regensburg
Regensburg, Germany
florian.habler@stud.uni-regensburg.de

Polina Ugnivenko
University of Regensburg
Regensburg, Germany
polina.ugnivenko@stud.uni-regensburg.de

Niels Henze
University of Regensburg
Regensburg, Germany
niels.henze@ur.de

Valentin Schwind
University of Regensburg
Regensburg, Germany
valentin.schwind@ur.de

CHI '20 Extended Abstracts, April 25–30, 2020, Honolulu, HI, USA.

ACM ISBN 978-1-4503-6819-3/20/04.
https://doi.org/10.1145/3334480.3382826

Abstract
Virtual reality (VR) enables immersive applications that make rich content available independent of time and space. By replacing or supplementing physical face-to-face meetings, VR could also radically change how we socially interact with others. Despite this potential, the effect of transferring physical collaborative experience into a virtual one is unclear. Therefore, we investigated the experience differences between a collaborative virtual environment (CVE) and a physical environment. We used a museum visit as a task since it is a typical social experience and a promising use case for VR. 48 participants experienced the task in real and virtual environments, either alone or with a partner. Despite the potential of CVEs, we found that being in a virtual environment has adverse effects on the experience which is reinforced by being in the environment with another person. Based on quantitative and qualitative results, we provide recommendations for the design of future multi-user virtual environments.

Author Keywords
Collaborative Virtual Reality; Virtual Reality; Virtual Avatar; Social Experience; Museum; Co-presence.

CCS Concepts
•Human-centered computing $\rightarrow$ User studies; Virtual reality;

Figure 1: The experiment's four conditions (from up to down): Two participants in virtual reality, a single participant in virtual reality, two participants in the physical space, and a single participant in the physical space.

Introduction & Background

Virtual reality (VR) enables experiences that can be similar to or completely different from the real world. Recent commercial development made VR affordable for the wider population. Consequently, VR is increasingly used in several fields, including entertainment, digital prototyping, therapy, training, and education [11, 12]. Collaborative Virtual Environments (CVEs) are especially an interesting domain of VR as they enable users to meet independently of their physical location virtually. CVEs were one of the first VR use cases [6, 23] and enable multiple users to collaborate and work in the same virtual environment [10]. As they reduce the need for traveling to share experiences, it has even been argued that CVEs could contribute to reduce our carbon footprint [18].

While CVEs do not require users to meet physically, and the VR technology has become widely available, research and industry still rely on physical meetings. Earlier work was mainly concerned with the technical challenges, such as the architecture [25] or the implementation [13] of CVE systems. Today, however, the number of massively multi-player online games, such as Rec Room[1] or the Tabletop Simulator for VR[2] are available and widely accessible. This shows that many technical challenges have been solved. Some reasons might explain why successful CVE applications are currently limited to the entertainment domain. For example, other users' avatars could trigger the uncanny valley effect [17, 22], a mismatch between a user's physical appearance and the user's avatar can be irritating [1], and imprecise tracking could demand users to be overcautious [3, 20]. The fidelity of CVEs dramatically increased in recent years. Thus, it is unclear how experiencing a CVE differs from experiencing a physical counterpart.

Investigating CVEs requires an environment that provides a believable benefit and social experience for multiple users. Visiting museums is a typical social activity. Several studies indicated the collaborative nature of museum visits by showing how couples or groups navigate, encounter, share and experience the museum items and how the visitors communicate their experience to other visitors or a museum guide [5, 24]. Previous work showed that museum visitors prefer to engage in social learning experiences [8]. These works indicate the significance of co-participating in museum experiences. Furthermore, cultural heritage is repeatedly proposed as a serious use case for VR [4]. VR enables virtual conservation and restoration of destroyed or damaged artefacts [2, 9, 14], as well as learning about cultural heritage [7, 15]. While virtual museums are a relevant research domain on their own, we focus on using them as an example of a social experience.

In this paper, we explore the differences between an immersive social virtual environment and a physical counterpart when being alone or together with a partner. We use museum experience as a social activity. We focus more on the accompanied visit than on the collaboration. In a study with 48 participants, we compared *virtual* and *real* museum environments. Participants experienced the real and the virtual museum either *alone* or *accompanied*. Our work showed that being in a virtual environment has adverse effects on the museum experience which is reinforced by another person. Based on quantitative and qualitative results, we provide recommendations to improve the design of future multi-user virtual environments.

[1] https://rec.net
[2] https://store.steampowered.com/app/286160/Tabletop_Simulator

Method

We conducted a study to learn how physical and virtual social experiences differ when being alone or accompanied. We simulated a museum visit as a typical social experience. Thus, we prepared an exhibition room in our lab that contained two exhibits placed at the opposite sides of the room. To counterbalance the order of exhibits, we prepared four exhibits. Each exhibit consisted of a bust and a painting of a composer. We placed the bust on a rectangle bar table that resembled a pedestal and attached the painting to a small whiteboard (see Figure 2). For each exhibit, we prepared an audio guide that described the life of the composer by referencing to the bust and the painting. It also contained a short music piece by the composer. The audio guides took on average 215 seconds. Thus, an exhibition scene contained two pairs of busts, paintings, and the audio guide. To recreate the exhibition in VR, we created a 3D model of the physical room and the exhibits (see Figure 3). For the VR conditions, participant's body was accordingly visualized using a high-precision motion capture system. Therefore, we used male and female avatars with two alternative versions by altering the color of their clothes (see Figures 1 and 3). Both the model and skeleton of the avatars were based on the Genesis 8 model in DAZ3D [3].

Figure 2: A pair of participants immersed in VR while standing physically as well as virtually in front of one of the exhibits.

We conducted a mixed-design study with two independent variables: VISITOR and ENVIRONMENT. Both independent variables had two levels resulting in four conditions (see Figure 1). We used VISITOR (*alone*, *accompanied*) as a between-subjects variable to reduce sequence effects. As the within-subjects variable, we used ENVIRONMENT (*virtual*, *real*) to enable participants compare the real and the virtual museums. As a dependent variable, we measured participants' comprehension of the exhibits. We prepared ten multiple-choice questions with four possible answers for each exhibit. The questions were based on the audio guides. The difficulty of the comprehension questions for each exhibit was similarly distributed and was in the acceptable range (30-70%) [16]. Furthermore, we used the Igroup Presence Questionnaire (*IPQ*) to determine the experienced presence [21]. For the *accompanied* conditions, we measured *co-presence* using the questionnaire by Poeschl and Doering [19]. Participants filled the questionnaires individually, even when experiencing the exhibition visits as a pair. Moreover, for the *accompanied* conditions, we recorded the participants' position to derive participants' relative distances to each other. Finally, participants were asked to provide qualitative feedback for each condition.

We recruited 48 participants (17 female) through our university's mailing lists. Their average age was 25 years (SD=3.8). We randomly distributed male and female participants across the two levels of VISITOR. However, we ensured that participants in the *accompanied* conditions knew each other. Participants were mainly students studying technical disciplines. 77.08% of participants had experience with VR.

As an apparatus, we used two HTC Vive headsets with wireless adapters to enable participants to move within the tracking volume freely. We used Unity 3D (Version 2018.3.11f1) to develop the virtual museum, and the scene was rendered by two identical high-performance PCs running Windows 10, Intel i7-8750H, 16GB RAM, and an NVIDIA GeForce GTX 1060 graphics card. To track participants in full-body motion, we used an OptiTrack motion tracking system[4] with twelve cameras (eight PRIME 13 and four PRIME 13W). The motion tracking software was running on a dedicated PC with Windows 10, Intel i7-8700, 26GB RAM, and an NVIDIA GeForce GTX 1080 graphics card.

[3] https://www.daz3d.com

[4] https://optitrack.com

Participants wore a marker-based full-body motion capture suit with 49 markers (see Figure 2). They could freely move within the 4.2m x 3.9m tracking volume. We marked the boundary of the tracking volume with white stripes on the floor, both in the real and virtual environments. Participants wore the full-body motion capture suits throughout all conditions to record their position in VR and the physical scenes. Independent of the environment, participants could activate the audio guides using the HTC Vive's controllers we handed them. The audio guides were played through speakers mounted on the ceiling.

Figure 3: One of the exhibits used in the study that consisted of a bust, a picture, and a verbal description that included an example of the composer's music. The upper image shows an exhibit from a female participant's perspective, and the lower image shows the exhibit from a third-person perspective with a female participant standing in front of it.

Depending on the condition, either one or two participants were invited to our lab. After introducing the aim and procedure of the study, participants signed a consent form and filled a demographic questionnaire. They were also asked to rate how important it is for a museum to have a virtual counterpart. Afterward, we introduced the HTC Vive and motion capturing suit to the participants and helped them to wear the suit. We adjusted the headset and calibrated the marker tracking system. During the calibration, participants were in a dedicated VR environment without the exhibition content. We told them to stand at the defined position in the middle of the room and ensured that all markers were recognized. For the conditions in the physical scene, participants were asked to take off the VR headset. We explained that the audio guide for each exhibit could be played only once by pointing at the paintings with the HTC Vive's controller while being close to it and pressing a trigger button. Participants were aware of the comprehension tests after each condition. Afterward, for both real-world and VR conditions, we asked participants to explore the exhibits within the tracking volume for ten minutes. We reminded participants how to activate the audio guides if they did not do it during the first five minutes. After ten minutes, we asked them to leave the tracking volume.

After exploring two exhibits, participants were asked to fill the comprehension tests about both composers and the *IPQ* questionnaire. Accompanied participants filled the *co-presence* questionnaire as well. After completing all questionnaires, participants continued with the second condition. At the end of the study, they were again asked to rate the importance of a museum to have a virtual counterpart and to give general feedback about the virtual and real-world museum visit experience.

Results

Quantitative Results

Each participant explored two exhibits in real and virtual museums, either alone or accompanied. For the evaluation, we performed a quantitative analysis of the collected objective and subjective data. To perform nonparametric tests using the two independent variables VISITOR and ENVIRONMENT, we used a linear mixed-model analysis using the ARTool package for R by Wobbrock *et al.* [26].

Comparing the *IPQ* scores, we found a significant effect of the ENVIRONMENT (see Table 3). While the sense of presence (see Table 1) were significantly lower in the *virtual* ($M = 4.03$, $SD = .72$) than in the *real* ($M = 4.62$, $SD = 1.05$) environment, we found no effect of being *alone* or *accompanied* on presence. *Co-presence* measures were obtained from participants in all *accompanied* conditions. Paired samples t-test revealed a significant difference between the *real* and *virtual* environments, $t(23) = 2.2186$, $p = .036$. The analysis shows that the scores were significantly lower when the participants experienced the scene in the *virtual* ($M = 4.425$, $SD = 1.147$) than in the *real* ($M = 4.997$, $SD = 1.111$) environment.

We asked participants before (*pre*) and after (*post*) the study to assess how important it is for a museum to have a virtual counterpart. We found no significant difference between pre- and post-experience, $F(1, 46) = 0.372$, $p = .544$, or of VISITOR, $F(1, 46) = 2.745$, $p = .104$. We found, however, a significant PRE-/POST × VISITOR interaction effect, $F(1, 46) = 4.953$, $p = .03$. The participants provided similar ratings before the experiment when *alone* ($M = 3.50$, $SD = 1.14$) and *accompanied* ($M = 3.33$, $SD = 1.52$). The ratings changed in opposite directions after experiencing the exhibitions *alone* ($M = 4.13$, $SD = 1.52$) and *accompanied* ($M = 3.00$, $SD = 1.53$). Thus, participants considered virtual museums more important after experiencing the exhibitions *alone* but less important after *accompanied* visits.

Regarding the number of correctly answered questions of the comprehension tests, we found no significant main or interaction effects (see Table 3). Table 2 summarizes the participants' comprehension results.

	M	SD
Real-Alone	4.79	1.06
Virtual-Alone	4.23	0.56
Real-Accompanied	4.46	1.04
Virtual-Accompanied	3.84	0.82

Table 1: *IPQ* presence scores for all conditions.

	M	SD
Real-Alone	11.75	2.56
Virtual-Alone	11.25	3.22
Real-Accompanied	12.25	2.98
Virtual-Accompanied	10.42	3.24

Table 2: Comprehension measures for all conditions.

To determine if participants behaved differently in the *virtual* and the *real* environments when *accompanied*, we measured the distance between pairs of participants. Paired samples t-test revealed a statistically significant difference between both levels, $t(11) = -2.705$, $p < .05$. The average distance between the participants was lower in the *real* ($M = 1.027$, $SD = .310$) than in the *virtual* ($M = 1.157$, $SD = .242$) environment.

Qualitative Feedback

At the end of the study, each participant provided feedback on the conditions. Participants were positive about the collaborative visit to museums (*e.g., "I did not like the feeling that I was alone. For well-being there should be more people in the museum."* (P1)) Interestingly, participants stated that in the *virtual* conditions they could better focus on the museum visit experience: *"... a good way to concentrate, simple environment, little distraction."* (P21) Since participants could freely walk in the *virtual* conditions, they also commented on the real to virtual world mapping: *"In the real museum, there were no hesitations to bump into something I did not see because of the VR-Glasses."* (P6) Consequently, participants indicated walking faster and closer to the exhibits in the *real* museum: *"[In relation to the virtual museum,] I walked closer to the busts."* (P4) *"[In the real museum,] I was watching the other person, and I walked around safer and faster."* (P27)

	ENVIRONMENT (df=1)		VISITOR (df=1)		ENVIRONMENT × VISITOR (df=1)	
	F	p	F	p	F	p
IPQ score	12.316	<.001	3.413	.071	.031	.86
Comprehension	3.358	.073	.068	.794	.735	.395

Error: df=46

Table 3: Results of the two mixed-model ANOVAs between ENVIRONMENT and VISITOR for IPQ and comprehension scores.

Since the museum contained only a few items to keep participants focus on the exhibits and the audio guides, eight participants stated that the *real* museum was *"simple and boring"*. However, no participant indicated the same for the *virtual* museum. Furthermore, during the *accompanied* visit to the *virtual* museum, participants were missing the mimic and eye contact of the other person: *"The partner was hard to read only by body language, facial expressions were missing."* (P40) P25 was distracted by the avatars: *"In VR, I was more distracted by my body and the other person than in the real world."* In comparison to the VR condition, participants indicated that they had more interaction with the other participant in the *real* museum: *"In the real world, I paid attention to the other person's facial expressions. That did not work in VR."* (P25)

Ten participants wanted to get informed not only through an audio guide but also by reading text labels in the museums: *"Learning-factor might increase with additional written information, not only with audio."* (P6) Participants also stated that a *virtual* environment should not be a replica of a *real* museum, and it should be possible to interact with the environment: *"The VR should enable something that is not possible in the real museum."* (P38) *"It would be great if I could zoom in and out the images. It should not be the same as in real life."* (P26)

Discussion

We conducted a study to investigate how social learning experience differs when taking place in real and virtual environments while simultaneously manipulating being alone or accompanied. We found that the virtual environment reduces the perceived presence independent of being alone or accompanied. When accompanied, we found that the virtual environment reduces co-presence and increases the interpersonal distance between persons. After experiencing the exhibits alone, participants considered virtual museums more important than before. Experiencing the exhibits accompanied, participants considered virtual museums less important than before.

The qualitative results revealed several explanations for the subjective results. We used high-precision motion tracking to embody participants in avatars. Thereby, we also showed their motion to the accompanying participant. This is, however, clearly not sufficient, as indicated by the lower co-presence and supported by qualitative comments. Facial expressions and eye contact are crucial for social interaction in the physical world and must also be conveyed in the virtual one. Independent of being alone or accompanied, the virtual environment caused a lower sense of presence. While participants' full-body motion was captured and displayed in VR, participants experienced a body different from their own. This resulted in being more cautious in the virtual environment. Especially participants experiencing the exhibits alone criticized that they indeed felt alone in the virtual environment.

Despite the quantitative results, participants' comments on the virtual environment and its potential were surprisingly encouraging. VR might enable users to focus better, as there can be even less distraction in VR. Virtual museums could provide experiences that are not possible in the real world but should not be replications of the physical ones. The virtual environment could offer possibilities to interact with content that are not possible in the physical world, such as zooming into the content. Some participants asked for content providing other modalities, such as textual labels. Future work should consider these suggestions.

Conclusion

In this paper, we investigated the differences between visiting a virtual and physical environment, either alone or accompanied. We found that replicating even a very simple physical environment while using high-quality motion capture results in a reduced sense of presence independent of being alone or accompanied. For accompanied experience, the virtual environment causes a decreased sense of co-presence and increased interpersonal distance. While as a virtual environment, we used an exact 3D model of the physical one without further interaction elements, future work should investigate virtual environments that embrace the possibilities that come with being in VR.

Acknowledgements

This work is supported by the German Ministry of Education and Research (BMBF) within the GEVAKUB project (01JKD1701A).

REFERENCES

[1] Jeremy N Bailenson, Kim Swinth, Crystal Hoyt, Susan Persky, Alex Dimov, and Jim Blascovich. 2005. The independent and interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments. *Presence: Teleoperators & Virtual Environments* 14, 4 (2005), 379–393.

[2] Fabio Bruno, Stefano Bruno, Giovanna De Sensi, Maria-Laura Luchi, Stefania Mancuso, and Maurizio Muzzupappa. 2010. From 3D reconstruction to virtual reality: A complete methodology for digital archaeological exhibition. *Journal of Cultural Heritage* 11, 1 (2010), 42–49.

[3] Marco A Bühler and Anouk Lamontagne. 2018. Circumvention of pedestrians while walking in virtual and physical environments. *IEEE transactions on neural systems and rehabilitation engineering* 26, 9 (2018), 1813–1822.

[4] Marcello Carrozzino and Massimo Bergamasco. 2010. Beyond virtual museums: Experiencing immersive virtual reality in real museums. *Journal of Cultural Heritage* 11, 4 (2010), 452–458.

[5] Luigina Ciolfi and Liam Bannon. 2002. Designing Interactive Museum Exhibits: Enhancing visitor curiosity through augmented artefacts. In *Eleventh European Conference on Cognitive Ergonomics*. 7.

[6] Christopher Codella, Reza Jalili, Lawrence Koved, J. Bryan Lewis, Daniel T. Ling, James S. Lipscomb, David A. Rabenhorst, Chu P. Wang, Alan Norton, Paula Sweeney, and et al. 1992. Interactive Simulation in a Multi-Person Virtual World. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '92)*. Association for Computing Machinery, New York, NY, USA, 329–334.

[7] Nicoletta Di Blas and Caterina Poggi. 2006. 3D for Cultural heritage and education: evaluating the impact. In *Museums and the Web 2006.* Archives and Museums Informatics, 141–150.

[8] John H Falk, Carol Scott, Lynn Dierking, Leonie Rennie, and Mika Cohen Jones. 2004. Interactives and visitor learning. *Curator: The Museum Journal* 47, 2 (2004), 171–198.

[9] Armin Grün, Fabio Remondino, and Li Zhang. 2004. Photogrammetric reconstruction of the great Buddha of Bamiyan, Afghanistan. *The Photogrammetric Record* 19, 107 (2004), 177–199.

[10] Olof Hagsand. 1996. Interactive multiuser VEs in the DIVE system. *IEEE multimedia* 3, 1 (1996), 30–39.

[11] Hsiu-Mei Huang, Ulrich Rauch, and Shu-Sheng Liaw. 2010. Investigating learners' attitudes toward virtual reality learning environments: Based on a constructivist approach. *Computers & Education* 55, 3 (2010), 1171–1182.

[12] Jason Jerald. 2015. *The VR book: Human-centered design for virtual reality*. Morgan & Claypool.

[13] Chris Joslin, Igor S Pandzic, and Nadia Magnenat Thalmann. 2003. Trends in networked collaborative virtual environments. *Computer Communications* 26, 5 (2003), 430–437.

[14] Marc Levoy, Kari Pulli, Brian Curless, Szymon Rusinkiewicz, David Koller, Lucas Pereira, Matt Ginzton, Sean Anderson, James Davis, Jeremy Ginsberg, and et al. 2000. The Digital Michelangelo Project: 3D Scanning of Large Statues. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '00)*. ACM Press/Addison-Wesley Publishing Co., USA, 131–144.

[15] Nadia Magnenat-Thalmann and George Papagiannakis. 2005. Virtual worlds and augmented reality in cultural heritage applications. *Recording, modeling and visualization of cultural heritage* (2005), 419–430.

[16] Gyata Mehta and Varsha Mokhasi. 2014. Item analysis of multiple choice questions-an assessment of the assessment tool. *International Journal of Health Sciences & Research* 4, 7 (2014), 197–202.

[17] Masahiro Mori, Karl F. MacDorman, and Norri Kageki. 2012. The Uncanny Valley [From the Field]. *IEEE Robotics Automation Magazine* 19, 2 (June 2012), 98–100.

[18] David M Pearlman and Nicholas A Gates. 2010. Hosting business meetings and special events in virtual worlds: a fad or the future?. In *Journal of Convention & Event Tourism*, Vol. 11. Taylor & Francis, 247–265.

[19] Sandra Poeschl-Guenther and Nicola Doering. 2015. Measuring Co-Presence and Social Presence in Virtual Environments - Psychometric Construction of a German Scale for a Fear of Public Speaking Scenario. *Annual Review of CyberTherapy and Telemedicine* 13 (01 2015), 58–63.

[20] Ferran Argelaguet Sanz, Anne-Hélène Olivier, Gerd Bruder, Julien Pettré, and Anatole Lécuyer. 2015. Virtual proxemics: Locomotion in the presence of obstacles in large immersive projection environments. In *2015 IEEE Virtual Reality (VR)*. IEEE, 75–80.

[21] Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. 2001. The experience of presence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments* 10, 3 (2001), 266–281.

[22] Valentin Schwind, Katrin Wolf, and Niels Henze. 2018. Avoiding the Uncanny Valley in Virtual Character Design. *Interactions* 25, 5 (Aug. 2018), 45–49.

[23] Haruo Takemura and Fumio Kishino. 1992. Cooperative Work Environment Using Virtual Workspace. In *Proceedings of the 1992 ACM Conference on Computer-Supported Cooperative Work (CSCW '92)*. Association for Computing Machinery, New York, NY, USA, 226–232.

[24] Dirk Vom Lehn, Christian Heath, and Jon Hindmarsh. 2001. Exhibiting interaction: Conduct and collaboration in museums and galleries. *Symbolic interaction* 24, 2 (2001), 189–216.

[25] Adrian West and Roger Hubbold. 2001. System challenges for collaborative virtual environments. In *collaborative virtual environments*. Springer, 43–54.

[26] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. Association for Computing Machinery, New York, NY, USA, 143–146.