Symmetric Evaluation: An Evaluation Protocol for Social VR Experiences

Nadine Wagener
University of Bremen
nwagener@uni-bremen.de

Johannes Schöning
University of Bremen
schoening@uni-bremen.de

Workshop on Social VR: A New Medium for Remote Communication Collaboration at CHI '20, April 26, 2020, Honolulu, HI, USA.

Abstract
Virtual reality (VR) technologies are nowadays widely used for remote collaboration. For example, they are used to design and evaluate product concepts with multiple users. While the design process is often fully conducted in VR, the evaluation hereof is mostly still performed outside the virtual environment (VE). Researchers have already started to study the effects of including evaluation methods in the VE, first and foremost the thinking-aloud protocol and in-VR-questionnaires. However, the experimenter typically stays outside the VE. In this workshop paper we outline why we believe that it is beneficial to *evaluate everything* within the VE and propose a fully symmetric in-VR evaluation protocol for social VR experiences.

Author Keywords
Symmetric Evaluation; Virtual Reality; Social Virtual Reality; Evaluation; Remote Collaboration

CCS Concepts
•**Human-centered computing** → **Virtual reality; Mixed / augmented reality;** *Empirical studies in HCI; User studies; Walkthrough evaluations;* Empirical studies in collaborative and social computing;

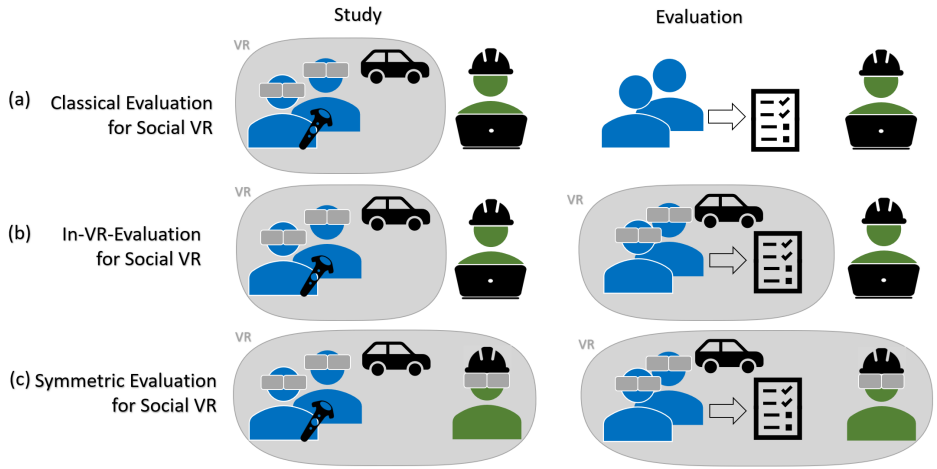

Figure 1: Study and evaluation in a Social VR setting: (a) Classical Structure: *study* - participants evaluate a product in VR while examiner is outside VR, *evaluation* - participants fill out a post-test survey, (b) In-VR-Evaluation: *study* - participants evaluate a product in VR while examiner is outside VR, *evaluation* - participants fill out a survey in VR while examiner is outside VR, (c) Symmetric Evaluation: *study* - participants evaluate a product in VR while examiner is in VR, *evaluation* - participants fill out a survey in VR while examiner is in VR

Introduction & Motivation

Virtual reality (VR) technologies are nowadays used across many industries to design products in a collaborative fashion. Already more than one third of US manufacturers plan to establish VR technologies in their production processes and it is estimated that these technologies will contribute about 1.5 trillion US$ to the global economy in 2030 [10]. Virtual environments (VEs) are used for evaluation purposes in order to get efficient and inexpensive feedback on visual and usability aspects of their product [3]. Karre et al. provide a good overview of how VR is used in the industry for evaluation purposes [8].

If the users are geographically dislocated, the industry relies on so called collaborative virtual environments (CVEs), in which two or more users can evaluate cooperatively at the same time in a VE [12].

Nevertheless, the actual evaluation step, including annotations, filling out questionnaires and giving interviews, is mostly performed outside VR as depicted in figure 1 (a) [11]. Thus, spontaneous thoughts and feedback are not reflected, which results in an inaccurate completion of the survey [5, 11]. Therefore, HCI researchers have already started to investigate the effects of different evaluation protocols in VR (see figure1(b)). Frommel et al. tested the concept of integrated questionnaires regarding computer

games with a positive outcome [5]. Haas found that in-VR-questionnaires lead to a higher truthfulness of the feedback compared to the thinking-aloud protocol, a method in which participants permanently express their thoughts out loud [7]. Based on these findings, Schwind et al. showed that presence questionnaires offer similar results when being completed in VR as compared to an evaluation done in reality [11]. Alexandrovsky et al. further verified the need for such an evaluation by conducting an expert survey in which 64% of the researchers stressed the importance of in-VR-questionnaires and 82% preferred the in-VR-questionnaire over a post-test evaluation [1].

However, in all these studies the experimenter remained outside the VE. We believe that it is beneficial to a study's quality that participants and the examiner experience the VE and the evaluation together. Thus, we propose a fully symmetric evaluation protocol for social VR experiences that we outline below (see figure1 (c)). We will also discuss important research questions when conducting future studies about symmetric evaluation that will offer thought-provoking impulses for the HCI-community.

Symmetric Evaluation for Social VR Experiences

Instead of either having participants filling out a questionnaire in VR or using the thinking-aloud protocol to interact with an *invisible experimenter*, in a symmetric evaluation the supervisor would be present in VR both for the duration of the study and the evaluation step. The examiner would keep his task of supervising and assisting the participants while the latter would fulfill their task of performing and evaluating.

We envision that one of the main advantages of participants and experimenter sharing a CVE is that some participants' feedback might only be correctly interpreted when having experienced the participants' behavior in the VE first-hand. This is especially true when exploring subjective facets as the feeling of 'being-there' called presence or the user experience. A symmetric evaluation should ideally also increase the participants' presence because they do not need to talk to seemingly invisible bystanders. In turn, a higher presence results in a better feedback of the product [4]. Furthermore, the supervisor can offer assistance when needed by pointing things out without complicated explanations and descriptions. As a symmetric in-VR evaluation is technically not different from any other multi-user setup, as described for example by Waldow et al. [13], it is also possible to provide view sharing and recording buttons as user interfaces for cooperatively editing and evaluating a VE [9]. These features enable co-workers to use collaborative problem solving (CPS) skills, which improve amongst others a shared understanding, team organisation and communicative cooperation in a self-regulated manner [6].

Future Work

Immersing one or more participants and study conductor both in VR at the same time opens up many different facets to explore. First, the effect on the participants of being able to interact with the supervisor in VR could be measured. In respect to that, the following research questions might be of special interest:

- How does having the experimenter in the VE affect the *behaviour* of the participant?
- How does having the experimenter in the VE affect the *performance* of the participant, e.g. task load, speed of task fulfillment, mistakes, amount of words spoken, questions asked?

- How does having the experimenter in the VE affect the *emotional state* of the participant?

We hypothesise that by visually sharing the VE with the supervisor the participants either feel assisted and more secure in their tasks or monitored and stressed. Perceiving assistance could result in, first, a better user experience and, second, less time needed for the evaluation procedure. If a feeling of being monitored arises, the participants' confidence could diminish, their stress level could increase and they could feel encouraged to answer in a socially desirable way.

We also foresee some serious decisions about the study's setup regarding the manner in which the experimenter can be included into the VE. This is tightly linked to research about avatar design, because embodiment forms the basis of social VR experiences [2].

- Should the experimenter have the possibility to engage with the participant at eye level, e.g. by having an own avatar?
- Should the avatar of the experimenter resemble themselves as much as possible or appear rather as an abstract humanoid model?
- How do these decisions about avatars affect the above mentioned research questions and the evaluation by the participant itself?

With today's technological advances, it is already possible for avatars to be equipped with a 3d facial scan of the research collaborators to make them look as realistic as possible [2]. On the one hand, participants of a study might find it easier to relate to this avatar as their assistant after having met the real instructor face-to-face beforehand. It further enhances the users' presence and simplifies social interaction [2]. On the other hand, though, when attempts to recreate human facets fall short, people might be affected by the Uncanny Valley Effect. This effect describes a rejection of a photorealistic avatar because it looks human but cannot (yet) convincingly mirror a real human being [2]. If this effect occurs, the participants might get distracted from their actual tasks and their distress with the avatar might negatively influence the evaluation itself.

A solution might be to include the examiner in an abstract way like a floating head-and-hands representation [2]. This approach still allows for an approximation of the supervisor's position while it might be less intrusive than full-body avatars [2]. No matter the manner of the supervisor's embodiment, it will most certainly influence the participants' behaviour, task result and emotional state in one way or another. If the impact is helpful or obstructive for the evaluation process itself should be researched.

Conclusion & Outlook

A symmetric evaluation protocol for social VR experiences is a multi-user VR setting in which one or more participants and the experimenter share the same VE at all times. This novel form of VR evaluation will probably enhance the participants' presence and lead to a better understanding of the given feedback. However, it offers space for debating the effect of a visible supervisor in the VE concerning the behavior, performance and emotional state of the users. It also results in design questions about how the examiner should be embodied that remain unanswered so far. We believe the HCI community plays an important role in illuminating these facets and we hope to stimulate discussions with our ideas in this workshop.

REFERENCES

[1] Dmitry Alexandrovsky, Susanne Putze, Michael Bonfert, Sebastian Höffner, Pitt Michelmann, Dirk Wenig, Rainer Malaka, and Jan D. Smeddinck. 2020. Examining Design Choices of Questionnaires in VR User Studies. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. DOI: http://dx.doi.org/10.1145/3313831.3376260

[2] Benjamin J. Bailey, Andrew Lilja, Campbell Strong, Katherine Moline, Maria Kavallaris, Rowan T. Hughes, and John McGhee. 2019. Multi-User Immersive Virtual Reality Prototype for Collaborative Visualization of Microscopy Image Data. In *The 17th International Conference on Virtual-Reality Continuum and Its Applications in Industry (VRCAI '19)*. Association for Computing Machinery, New York, NY, USA, Article Article 66, 2 pages. DOI: http://dx.doi.org/10.1145/3359997.3365730

[3] Leif P. Berg and Judy M. Vance. 2017. Industry use of virtual reality in product design and manufacturing: a survey. (2017). DOI: http://dx.doi.org/10.1007/s10055-016-0293-9

[4] Jennifer Brade, Mario Lorenz, Marc Busch, Niels Hammer, Manfred Tscheligi, and Philipp Klimant. 2017. Being There Again Presence in Real and Virtual Environments and Its Relation to Usability and User Experience Using a Mobile Navigation Task. *Int. J. Hum.-Comput. Stud.* 101, C (May 2017), 76–87. DOI: http://dx.doi.org/10.1016/j.ijhcs.2017.01.004

[5] Julian Frommel, Katja Rogers, Julia Brich, Daniel Besserer, Leonard Bradatsch, Isabel Ortinau, Ramona Schabenberger, Valentin Riemer, Claudia Schrader, and Michael Weber. 2015. Integrated Questionnaires: Maintaining Presence in Game Environments for Self-Reported Data Acquisition. DOI: http://dx.doi.org/10.1145/2793107.2793130

[6] Arthur C. Grasser, Peter W. Foltz, Yigal Rosen, David Williamson Shaffer, Carol Forsyth, and Mae-Lynn Germany. 2018. Challenges of Assessing Collaborative Problem Solving. In *Assessment and Teaching of 21st Century Skills: Research and Applications*, Esther Care, Patrick Griffin, and Mark Wilson (Eds.). Springer International Publishing AG, Switzerland, Chapter 5, 75–89. DOI: http://dx.doi.org/10.1007/978-3-319-65368-6_5

[7] Nico Haas. 2017. Evaluation of 'In-VR-Questionnaires'. Universität Stuttgart. DOI: http://dx.doi.org/10.18419/opus-10022

[8] Sai Anirudh Karre, Neeraj Mathur, and Y. Raghu Reddy. 2019. Usability Evaluation of VR Products in Industry: A Systematic Literature Review. In *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing (SAC '19)*. Association for Computing Machinery, New York, NY, USA, 1845–1851. DOI: http://dx.doi.org/10.1145/3297280.3297462

[9] Cuong Nguyen, Stephen DiVerdi, Aaron Hertzmann, and Feng Liu. 2017. CollaVR: Collaborative In-Headset Review for VR Video. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST '17)*. Association for Computing Machinery, New York, NY, USA, 267–277. DOI:http://dx.doi.org/10.1145/3126594.3126659

[10] PricewaterhouseCoopers (PwC). 2019. Seeing is believing: How VR and AR will transform business and the economy? (2019). Retrieved on January 30, 2020 from https://cloud.uk.info.pwc.com/seeing-is-believing-report-download.

[11] Valentin Schwind, Pascal Knierim, Nico Haas, and Niels Henze. 2019. Using Presence Questionnaires in Virtual Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, Article Paper 360, 12 pages. DOI: http://dx.doi.org/10.1145/3290605.3300590

[12] Misha Sra, Ken Perlin, Luiz Velho, and Mark Bolas. 2018. Novel Interaction Techniques for Collaboration in VR. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems (CHI EA '18)*. Association for Computing Machinery, New York, NY, USA, Article Paper W33, 8 pages. DOI: http://dx.doi.org/10.1145/3170427.3170628

[13] Kristoffer Waldow and Arnulph Fuhrmann. 2019. Using MQTT for Platform Independent Remote Mixed Reality Collaboration. In *Mensch und Computer 2019 - Workshopband*. Gesellschaft für Informatik e.V., Bonn. DOI:http://dx.doi.org/10.18420/muc2019-ws-570