innteract conference

Technische Universität Chemnitz
Fakultät für Maschinenbau
Professur Arbeitswissenschaft
und Innovationsmanagement

Using Virtual Prototyping Technologies to Evaluate Human-Machine-Interaction Concepts

Jennifer Brade [1], Manuel Dudczig [2], Philipp Klimant [3],

[1] jennifer.brade@mb.tu-chemnitz.de

[2] manuel.dudczig@mb.tu-chemnitz.de

[3] philipp.klimant@mb.tu-chemnitz.de

[1,2,3] Professorship for Machine Tools and Forming Technology,
Chemnitz University of Technology, Germany

DOI: 10.14464/aw&i conference.v3i0.241

ABSTRACT

Innovative products are interacting with the users based on smart sensors and algorithms. Logistics as one example is changing when automated guided vehicles are integrated in the process, supporting or even replacing workers. The way these products are interacting with humans and how they react to certain situations, will determine usability and user experience and therefore the success of use. Developing such products is based on innovative concepts that need to be evaluated and refined at early project stages. Using virtual-reality based user scenarios is one adequate option to do so. This paper describes technical as well as study-based approaches on how potential concepts are realized as virtual prototypes and evaluated by real users. It concludes with the evaluation results of a pilot study but also with general limitations and benefits as best practice advice for this kind of virtual prototyping techniques.

Keywords: *Virtual Prototyping, User Studies, User Experience, Virtual Reality, Logistics*

1 INTRODUCTION

Creating new products in general is becoming more complex in terms of mechanical and electrical parts that are used as well as lines of code that are written. Developing innovative human-centered products is a matter of involving future users and stakeholders within the early stages of the product development. Being innovative is about testing and evolving new concepts of interaction between the user and the product and benchmarking their potential. In early project stages this is usually done by using virtual and physical prototypes instead of completely realized solutions. One option to try new ways of human-machine interactions is using virtual reality (VR) within user tests. Besides necessary physical prototypes for subsequent validations this technology offers freedom of design for the product as well as the interactions. Using virtual prototypes within VR is a promising approach to speed up the interaction design process due to early user feedback of new concepts before physical solutions are available for user tests (Aromaa et al., 2014).

The framework of the described research is provided by a project called FOLLOWme. Main goal is the product development of an automated guided vehicle (AGV) for intralogistics that follows and interacts with the user helping to locate articles quicker and to reduce the error rate within picking processes. The AGV is developed together with industrial partners in the field of logistics, sensors and 3D-technologies to assure innovative and user-centered functionality and interaction design. The focus of this paper is the early VR-based evaluation of developed interaction concepts between the user and the AGV in user studies.

2 THEORY

Before virtual prototypes can be used to assess the user interaction of new concepts, the relevant factors for this assessment have to be defined. For a general understanding of the term virtual prototypes the definition of Wang (2002) can be used: *"A virtual prototype, or digital mockup, is a computer simulation of a physical product that can be presented, analyzed and tested by concerned product life cycle aspects such as design/engineering, manufacturing, service, and recycling as if a real physical model. The construction and testing of a virtual prototype is called virtual prototyping"* (p. 232). Even a virtual prototype cannot fully cover every aspect of a real product, nevertheless it is a valid alternative to a physical prototype, because of the advantages concerning time and cost savings in the product development process. But when using a virtual prototype instead of a real product, it is also important to know which effect the virtual world has on the user and the evaluations. Several studies have researched the differences between real and virtual environments in general and for specific user centered factors. An essential aspect of virtual environments is the factor presence, which is described as the *"sense of being in the virtual environment"* and is seen as a cognitive state that results from information processing of stimuli in the environment from various senses (Slater & Wilbur, 1997). Studies that compared real and virtual environments concerning presence show, that there are significant differences between the environments, but also indicate that when these differences are considered in the assessment of the test, a virtual scenario is a good alternative to a real setup. The study by Usoh et al. (2000) which was extended by Nisenfeld (2003) compared an environment shown with a head-mounted display (HMD) with a real environment. The studies by Busch et al. (2014) and Brade et al. (2017) compared a five-sided CAVE with a real laboratory or rather a real field environment concerning presence, which was measured with the ITC-SOPI (Lessister et al., 2001). All three studies show, that there are significant higher values for the factor negative effects (adverse psychological reactions) in the virtual environment and that the real environment exhibit significant higher values for the factor ecological validity (believability and lifelikeness of the scene). These differences have to be carefully considered within virtual studies.

For the assessment of products, whether in a real or a virtual environment, user centered factors like usability and user experience (UX) play an essential role. Usability describes the fitness of use of a product and *"extend to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use"* (Litwinowicz & Williams, 1994). The factor UX also measures usability factors like efficiency, dependability, fault tolerance, learnability and effectiveness, but extends these with factors like aesthetics, joy-of-use and attractiveness as Rauschenberger et al. (2013) remark. To properly evaluate the results of a usability or UX assessment of a virtual prototype, it is also important to keep in mind that the evaluation can be different between real and virtual environments. Brade et al. (2017), for example, found significant higher values for usability factors in the real environment and significant higher values for non-task oriented factors for the virtual environment in their study. The novelty of virtual reality itself and consequential the positive or negative influences on the ratings of these factors should be observed, when using virtual environments or virtual prototypes.

Beside these general studies in virtual environments or with virtual prototypes there are only a few studies which address picking scenarios in virtual environments. One study which compares different picking techniques is the study by Reif et al. (2007), where they opposed four picking techniques (picking list, VR-picking, pick-by-voice and AR-picking) in a within-subject design with 17 participants who had to solve five picking orders in each technique. The results showed, that the motivation of the participants was noticeable higher in the VR- and AR-based picking task. Furthermore they measured the cognitive load of the different techniques and showed that the pick-by-voice tasks cause the highest cognitive load. They also found that the cognitive load of the VR-scenario is higher than the real picking list, but not so high, that the virtual navigation and interaction has a high effect on the cognitive load.

The given overview of relevant concepts and studies of virtual prototypes shows, that there are some limitations and effects in the use of virtual prototypes compared to real prototypes. Table 1 gives a concentrated overview of advantages and disadvantages to be considered.

Table 1: Characterization of the study sample
Source: Own representation

Advantages of VR	Disadvantages of VR
• High experimental control • Higher motivation of the participants • Higher values of non-task oriented factors (hedonic quality aspects) • Easier and harmless feasibility	• Negative effects (adverse psychological reactions) • Cognitive load is higher • Usability has to be high, because participants are more critical • A high ecological validity is necessary, therefore high effort to create a VR scenario
Conclusion:	

VR-scenario needs to be as realistic as possible, positive effects of VR (motivation etc.) has to be noted, when VR results are compared with tests in reality

2.1 EVALUATED INTERACTION CONCEPTS

The standard EN ISO 9241-220 (2017) refers to four components for the quality of human-centered products: usability, accessibility, user experience and avoidance of harm. Within a human-centered product development all of these components need to be considered. The illustrated VR-based user tests mainly cover the demands of usability and user experience. Accessibility and avoidance of harm as factors are not described although VR as an interactive technology for user tests is able to cover those issues during product development phases as well.

The goal for the human-centered design of the AGV within the context of intralogistics is an efficient, ergonomic commissioning process with zero failure rates. Beside these specific project aims an increased usability will have competitive advantages. Making products easier to understand, reducing unsafe behavior due to a lack of usability and improving the user`s experience are further benefits of a high usability.

Looking at the AGV design, following actions belong to the commissioning process and therefore need to be designed as user friendly human-machine interactions within the virtual prototype to gain early user feedback:

- log on/off the AGV
- pick and place articles into boxes of the AGV
- correction of missing or misplaced articles
- status indication (AGV status and picking status)
- manual operation (navigate or replace the AGV)

To develop a successful solution for these goals and actions, several technical solutions were investigated. These options were combined to variable concepts that again were benchmarked considering weighted criteria using a preference matrix processed by project partners in a meeting to rate possible interaction concepts. Following three concepts were rated highest and therefore virtually realized and tested for user feedback and concept refinement before its technical implementation:

- 1. smart gloves including a scanner and use of a smart watch
- 2. smart gloves including a scanner and use of augmented-reality glasses
- 3. smart gloves including a scanner and use of text-to-speech output

The stated concepts were realized within a virtual intralogistics setting with common picking tasks to empower study participants to evaluate the interaction concepts close to the real world demands. The user receives the necessary picking information (article, amount, position as hallway, rack, shelf and box number) for each picking task visually on the smart watch, augmented-reality glasses or aural via a text-to-speech output. After locating the articles the user scans it with the smart glove with integrated scanning function and puts it in the transport box for transfer. To adjust the ergonomics of the AGV in all three scenarios it is possible to set the height of the AGV-platform and to manually adjust the position of the AGV in accordance to the racks via buttons and slider at the AGV input panel. The workflow and technical setup for this virtual reality setting to investigate the quality of human-machine interactions, ergonomic issues and the overall product perception are illustrated as follows.

2.2 WORKFLOW AND TECHNICAL SETUP

Depending on the type of product and the complexity of human-machine interactions the virtual prototype will vary in effort and implementation. Based on product or service ideas and possible technical or interaction concepts the context of constraints and field of application needs to be considered first. Once the setting is defined, a virtual reality scenario needs 3D objects that might deviate from computer-aided engineering software (CAx-Data) (often still in designing process) or computer generated imagery (CGI) assets that often reflect the operational environment (e.g. machinery, building, picking items). To gain a realistic impression even within a virtual setting, enhanced (e.g. physical based, texturized) materials, object animations and planned interactions (based on object colliders, physics and behavior scripts) are added to realize a close to reality system behavior. The technical implementations (see Fig. 1) for VR scenarios are often realized with game engines (Unity3D [13] in this case). In contrast to available visualization tools that are suitable for pictures and pre-rendered videos, a real-time rendering software is necessary for displaying interactive VR-prototypes. This type of software offers functionalities that are demanded in up-to-date 3D games as well as hardware support for HMDs to create stereoscopic, scale 1:1 scenarios to rate the usability and user experience in the user tests (Allmacher et al., 2018).

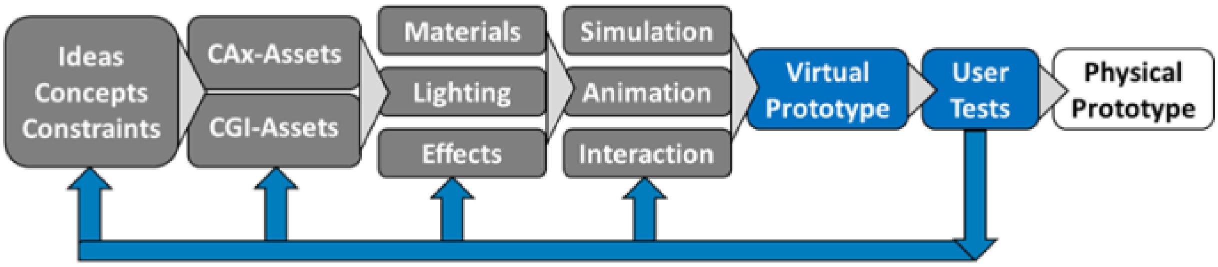

Figure 1: Project-Based Workflow of Virtual Prototypes
Source: Own representation according to [14]

In addition to the described workflow the used technical setup is illustrated in Figure 2. A high performance render PC creates the real-time user insight to the virtual world using a HMD (e.g. HTC Vive). The HMD is featured with an inside-out tracking based on non-visible rays. The HMD and two hand controllers are tracked within a room-scale area of up to 3x4m for high immersion. In contrast to the use of digital human models in virtual prototypes, the future user is interacting with the virtually mapped interactive system. User actions like walking physically, stretching your arms towards objects, bending down and so forth are necessary to properly evaluate interactions and ergonomic issues comparing different technical solutions to detect critical issues but also benefits (Allmacher et al., 2018).

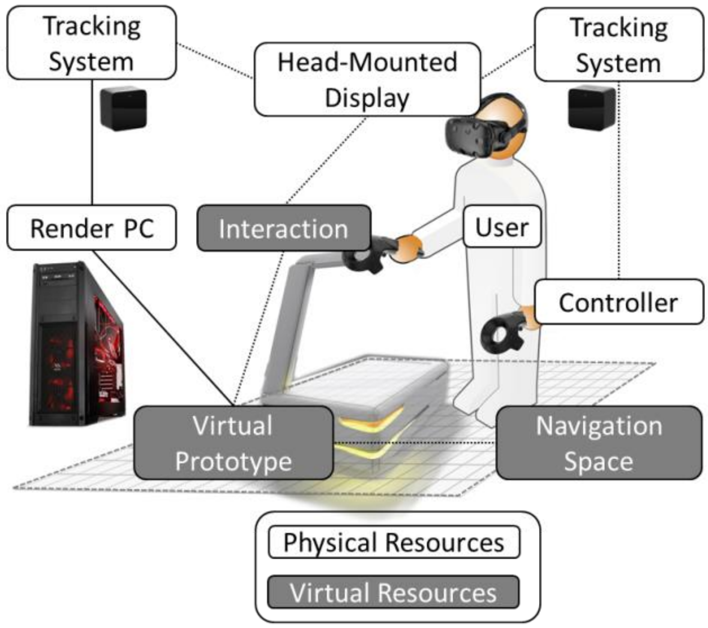

Figure 2 Project-Based Technical Setup
Source: Own representation according to Allmacher et al., 2018

Within the virtual scene static and dynamic objects and their instances need to be connected and triggered at certain events (e.g. picking up an object, scanning articles, provoking failures). Therefore the scene design needs to reflect the desired interactions of the technical concepts (illustrated in Fig. 3).

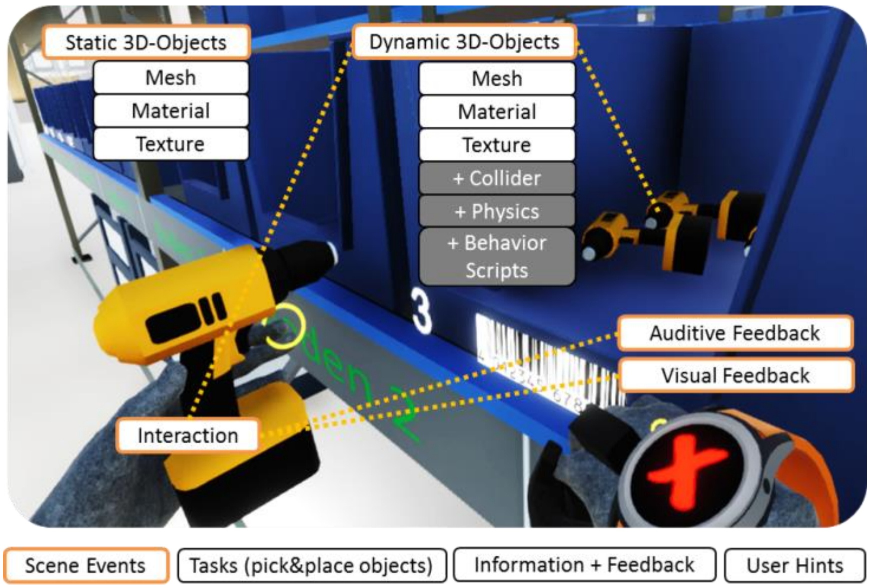

Figure 3: Project-Based Scene Structure for VR User Tests
Source: Own representation according to Allmacher et al., 2018

It also contains comprehensive and intuitive instructions for the user to minimize distortion of results due to variable explanations of the study supervisor. User-machine or user-object interactions are commonly based on collision detection of virtual objects (e.g. virtual hand and pick item) and should contain physical adequate behavior to not confuse the user according to his/her real life experiences and understanding of systems and behavior.

3 STUDY FRAMEWORK

To identify which of the developed interaction methods is preferred by the users, a pilot study was conducted to compare the interaction methods in a virtual picking-scenario. Therefore a user study as a within-subject design, with the different interaction methods as independent variable was carried out. In the study seven dependent variables using post-test questionnaires were collected: six user experience factors (attractiveness, perspicuity, efficiency, dependability, stimulation and novelty), measured with the User Experience Questionnaire (UEQ), and usability, measured with the System Usability Scale (SUS). The UEQ by Laugwitz et al. (2018) consists of 26 bipolar items that are assessed by a seven-point semantic differential (-3 to 3). Three scales (perspicuity, efficiency and dependability) describe the pragmatic quality of the tested methods, whereas two scales (stimulation and novelty) identify the hedonic quality. The scale attractiveness determines the general attitude to the tested method referring to Rauschenberger et al.(2013). Whereas the hedonic quality (non-task oriented quality aspects) is only determined by the UEQ, for the pragmatic quality (task oriented quality aspects), the SUS was used in addition. The ten-item SUS by Brooke (1996) rates the usability on a five-point Likert scale. Further a questionnaire was used to collect the demographic variables: age, gender, education and experience with virtual reality and picking processes.

The study itself consists of three parts: introduction and tutorial, main part and post-assessment. Each participant completed all three parts of the study. In the first part of the study the participants read the general instruction, which contained the explanation of the study aims and a declaration of the study tasks. Afterwards the participants put on the HMD and took the controllers given by the instructor. Then the instructor started a short tutorial, so that the users became familiar with the general interactions of the HTC Vive (used HMD). After completing the tutorial, the main part of the study started. It consisted of three picking orders, each one with another interaction method: The first scenario (see Fig. 4) is leaned on the interaction with a smartwatch. The virtual smartwatch was positioned on one hand and was combined with a scan glove. All picking orders were displayed on the smartwatch display. In the second scenario (see Fig. 4) the participants received the picking information through an augmented reality display in terms of canvas and different panels which were displayed in the field of view of the user. In the third scenario (see Fig. 4), the participants received the information through a realistic voice output.

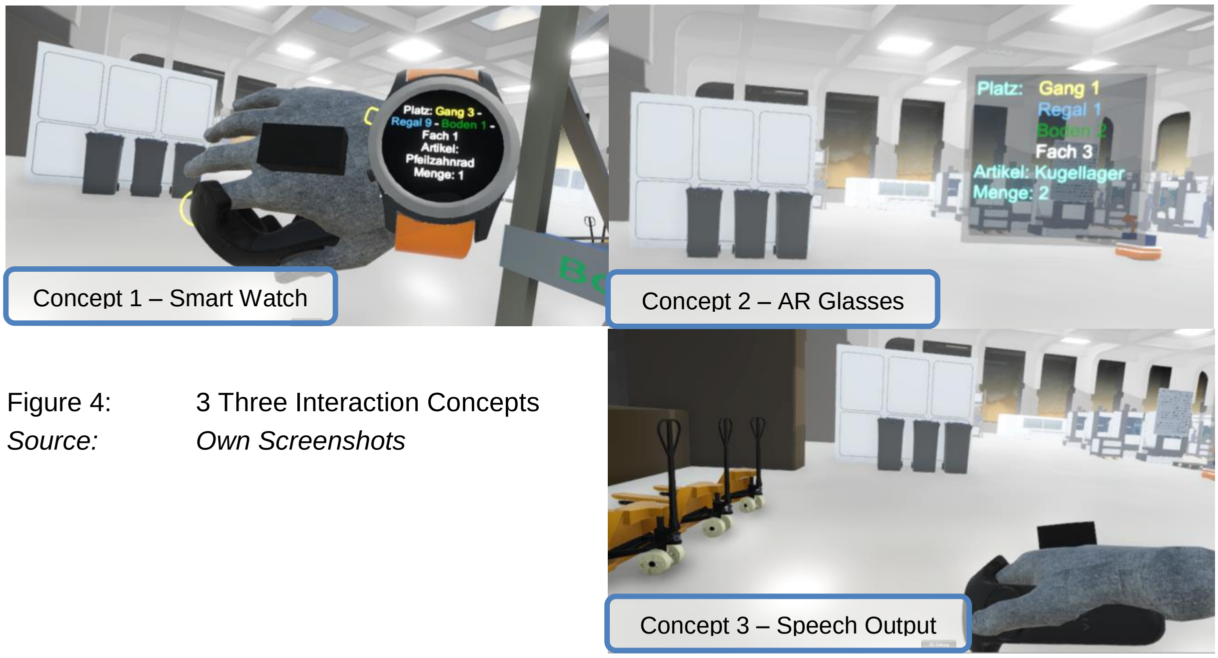

Figure 4: 3 Three Interaction Concepts
Source: Own Screenshots

All orders and information that are necessary, (e.g. choosing a wrong article) were presented by the voice output. During the picking order, the participants could play the instructions again and again by clicking a button on the controller. In all three scenarios, the picking orders are presented as equal as possible, to reduce effects of the representation.

Each picking order contains the necessary information (product name, number of products and the detailed position of the product) that the participant obtains through the specific interaction method. Moving in the virtual scenario is realized via a teleportation function. Through the task an interaction with the AGV was not necessary (unless for ergonomic reasons like repositioning of the AGV). When the picking order is completed the next order appears and the sequence was repeated.

To avoid sequence effects the order in which the interaction methods were tested was randomized beforehand. After finishing all picking orders in the main part, the participants received the printed questionnaire which assessed the demographic variables. Deviations and remarks that were made by the participants were noted in a study protocol and also confounding variables were registered. All participants received a small financial compensation or course credits for their participation.

Overall 10 participants, all students of the Chemnitz University of Technology, attended the pilot study. In Table 2 the sample is characterized in detail.

Table 2: Characterization of the study sample

Source: Own representation

Age (mean)	24,7 (SD=2,11)
Gender (% female)	40%
Previous contact with VR-Systems (% yes)	50%
Previous contact with HMDs (% yes)	30%
Previous contact with picking tasks (% yes)	10%

The evaluation of the SUS scores (see Fig. 5) shows a high usability rating for all three concepts. Based on the interpretation of the SUS score by Bangor et al. [18], the usability values for concept 1 – Smart Watch – (83,5%, SD=13,90), concept 2 – AR Glasses – (82,25%; SD=12,72) and concept 3 – Speech Output – (75,5, SD=15,67) indicate a good usability. The values of the user experience factors between the concepts are also shown in Figure 5.

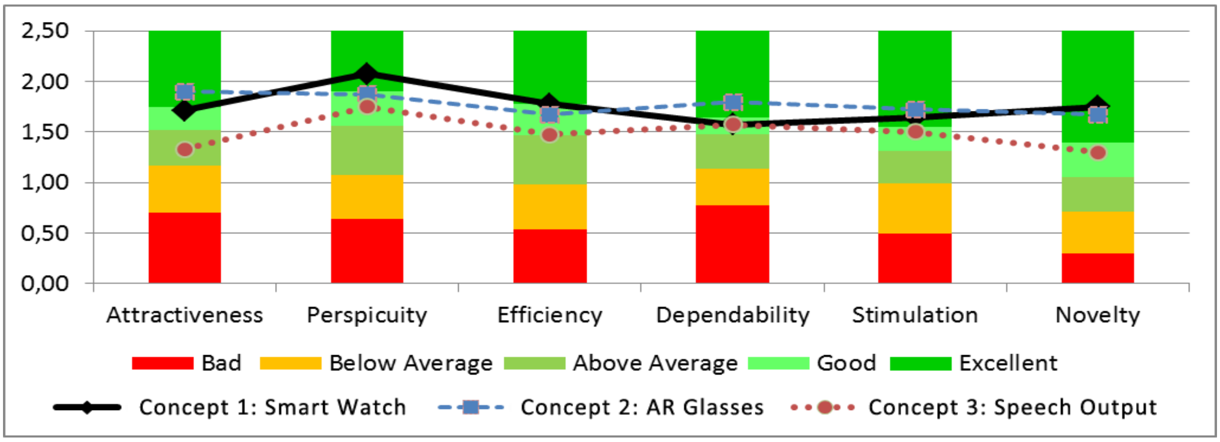

Figure 5: Benchmark of the User Experience Questionnaire

Source: Representation using template from www.ueq-online.org

Overall the means of three out of six factors (perspicuity, efficiency and novelty) are higher for the Smart Watch concept. On the other hand the means of the attractiveness, dependability and stimulation scales are higher for the AR-glasses concept. The text-to-speech concept reaches in all six factors the lowest scores of all three concepts.

These results show that the smartwatch concept reaches the highest usability score in the SUS, as well as in two out of three factors of the pragmatic quality aspects in the UEQ. In contrast the AR glasses concept reaches the highest values for the scale attractiveness, which means, the general attitude to the smart glass concept is more positive than to the other concepts. Nevertheless these differences between the concepts are only low and need to be verified in a study with more participants, preferably future system users. Because of the general high ratings of the UEQ and SUS factors for all three concepts, all three developed techniques are suitable for more tests, refinements and for the implementation in a real picking scenario.

The feedback of the participants was overall positive and the enjoyed the experience. The study was experienced as very interesting but some criticized that the voice output was very long and therefore partially annoying.

4 LESSONS LEARNED – BENEFITS AND LIMITATIONS

Using virtual reality prototypes in user tests to design and evaluate human-machine interaction in several projects revealed following basic *crucial factors for successful tests*:

- usability of VR-scenario is essential (using tutorials of how to use features, how to navigate and how to manipulate objects)
- intuitive interactions with virtual objects (convenient navigation, self-explained interactions, animated hints)
- multi-sensory feedback (visual, aural, haptic)

From the experience of the described and prior VR-based user tests there a few *limitations* that are currently hard to overcome:

- tracking controllers instead of real hand gestures to interact (approach: change of input devices and development of individual and stable interpreted gestures)
- existing influence of virtual prototypes leads to slightly deviating results compared to tests with physical prototypes (approach: final validations with physical prototypes)
- strongly varying experience of the test participants regarding Virtual Reality or the real life use case will have an significant effect on the results (approach: a higher number of test participants will balance this effect)

Following technical factors dependent on the best available VR technology are restricting the quality of the virtual prototypes:

- vast effort for realistic scenes (approach: automated conversion and reduction of CAx-data, using templates and assets)
- adequate render framerate while using complex, realistic and highly interactive scenes (approach: further scene optimization, hardware accelerations)
- quality of HMD-output (framerate > 60 fps, low latency from tracking system, smooth animations and transitions at ~ 60fps)
- restricted picture resolution within HMD (approach: user based HMD adjustment, future hardware evolution, e.g. using HTC Vive Pro instead)

Besides the crucial factors and limitations that need to be considered creating VR-based user tests there are many *benefits* that are worth the extra effort:

- saving development time and reducing the amount of physical prototypes

- testing concepts in iterative loops
- developing interface prototypes and interaction possibilities early to evaluate innovative ideas
- ergonomic tests with real people
- gaining early user feedback for concepts and product development

For future tests there are some additional factors that should be observed:

- compare the study with the results of other studies in VR, the factors presence, technical affinity, motivation and cognitive load should be collected
- beside the test in VR it is necessary to conduct the same test in reality to compare the results with recent studies and future studies
- a separate study to find out, which is the ideal length and kind of the voice output is necessary
- the quality of the optical information needs to be compared and should be as equal as possible (background color and form of the displays)

5 CONCLUSION AND OUTLOOK

This paper emphasizes the importance of early concept prototypes for developing, evaluating and refining innovative human-machine interactions. How virtual reality technologies can be used to realize interactive prototypes is described in terms of the workflow, scene and technical setup. To gain significant results, an adequate study setup and appropriate questionnaires need to be considered. The paper gives a practical example of a concept benchmark for interactions between users and automated guided vehicles within the context of logistics. The results of a pilot study, evaluating three interaction concepts, indicate the deliverables that will be essential for decision and process refinement during the interaction development. The paper closes with lessons learned that identify crucial factors, limitations but also benefits of using virtual reality prototypes for concept evaluations of human-machine interactions with user studies.

To increase the quality and relevance of virtual reality based concept studies, the render process as well as the hardware of HMDs and tracking systems still needs to be evolved for more realism of virtual prototypes. Using hands instead of controllers and collaborating with other users in real-time will improve the intuitive use of VR test studies additionally. Nevertheless, physical prototypes will still keep their eligibility for final validation of the product behavior and evaluation of all aspects of an interaction concept.

6 ACKNOWLEDGEMENTS

The research project “FOLLOWme ILS – Intralogistics System using Automated Guided Vehicles” has been supported by the German Federal Ministry of Education and Research.

REFERENCES

Aromaa, S.; Leino, S.-P.; Viitaniemi, J. (2014): Virtual Prototyping in Human-Machine Interaction Design. VTT Technology.

Wang, G. G. (2002): Definition and Review of Virtual Prototyping 2, pp. 232

Slater, M.; Wilbur, S. (1997): A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments, Presence: Teleoperators and Virtual Environments 6 (6), pp. 603-616

Usoh, M.; Catena, E., Arman; S., Slater, M. (2000): Using Presence Questionnaires in Reality, Presence: Teleoperators and Virtual Environments 9 (5), pp. 497-503

Nisenfeld, S. (2003): Using Reality to Evaluate the ITC Presence, Unpublished Master Thesis.

Busch, M.; Lorenz, M.; Tscheligi, M.; Hochleitner, C.; Schulz, T. (2014): Being There For Real – Presence in Real and Virtual Environments and its Relation to Usability, NordiCHI.

Brade, J.; Lorenz, M.; Busch, M.; Hammer, N.; Tscheligi, M.; Klimant, P. (2017): Being There Again – Presence in Real and Virtual Environments and its Relation to Usability and User Experience: International Journal of Human-Computer Studies 101, pp. 76-87

Lessister, J.; Freeman, J.; Keogh, E.; Davidoff, J. (2001): A Cross-Media Presence Questionnaire: The ITC-Sense of Presence Inventory, Presence: Teleoperators and Virtual Environments 10 (3), pp. 282-297

Litwinowicz, P.; Williams, L. (1994): Animating images with drawings. In Andrew Glassner, editor, *Proceedings of SIGGRAPH ’94* (Orlando, Florida, July 24–29, 1994), Computer Graphics Proceedings, Annual Conference Series, pages 409–412. ACM SIGGRAPH, ACM Press, July 1994.

Rauschenberger, M.; Schrepp, M.; Perez-Cota, M. et al. (2013): Efficient Measurement of the User Experience of Interactive Products. How to use the User Experience Questionnaire (UEQ).Example: Spanish Language Version. International Journal of Artificial Intelligence and Interactive Multimedia, Vol. 2, no. 1, pp. 39-45

Reif, R. et al. (Ed.) (2007): Einsatz von Virtual und Augmented Reality: Studie zur menschintegrierten Simulation und Prozessunterstützung im logistischen Umfeld. Bayerischer Forschungsverbund Supra-adaptive Logistiksysteme, Garching, Germany. Online: http://www.fml.mw.tum.de/PDF/ForLog_Studie_VR-AR_2007_Abgabe.pdf.

Standard prEN ISO 9241-220:2017, Ergonomics of human-system interaction - Part 220: Processes for enabling, executing and assessing human-centred design within organizations (ISO/DIS 9241-220.2:2017).

Unity3D. Unity Technologies. https://unity3d.com.

Allmacher, C.; Dudczig, M.; Klimant, P.; Putz, M. (2018). Virtual Prototyping Technologies Enabling Resource-Efficient and Human-Centered Product Development. Procedia Manufacturing, 21, pp. 749-756

Laugwitz, B.; Held, T.; Schrepp, M. (2008): Construction and Evaluation of a User Experience Questionnaire, HCI and Usability for Education and Work, Springer-Verlag Berlin Heidelberg Graz, pp. 63-76

Rauschenberger, M.; Schrepp, M.; Perez-Cota, M.; Olschner, S.; Thomaschewski, J. (2013): Efficient Measurement of the User Experience of Interactive Products. How to use the User Experience Questionnaire (UEQ).Example: Spanish Language Version, International Journal of Interactive Multimedia and Artificial Intelligence 2 (1) pp. 39-45

Brooke, J. (1996): SUS - A quick and dirty usability scale, Usability Evaluation in Industry, Taylor & Francis Inc. Bristol, pp. 189-194

Bangor, A., Kortum, P., & Miller, J. (2009). Determining what individual SUS scores mean: Adding an adjective rating scale. Journal of usability studies, 4(3), pp. 114-123