Mixed Reality in Asymmetric Collaborative Environments: A Research Prototype for Virtual City Tours

Nico Feld*
University of Trier

Benjamin Weyers†
University of Trier

Abstract

Augmented and virtual reality have the potential to improve various aspects of everyday life such as tourism. An elementary part of tourism is the organization and execution of guided city tours. In case of tours for virtual cities, such as in the case of no longer existing ancient settlements, a combination of augmented and virtual reality may render high benefit for collaboration between guide and visitors. Immersive technology has been shown suitable for collaborative systems by previous research, as they can provide a better experience than other technologies. However, recent literature analyses showed that there is a lack of research regarding the role of augmented reality supporting the active/leading user role in such scenarios. Usually, augmented reality is used for the users who are passively experiencing the virtual environment and not for the one controlling it. However, we think that augmented reality has a large potential to improve the workflow of such a primary user. To investigate this research question, we present a collaborative system in this work, which implements a virtual city tour. In this system, the guide is using augmented reality and the visitor is immersed in the virtual city using virtual reality glasses. To implement a valid research prototype, the system has been developed in close feedback with experts in the field of guided city tours.

Index Terms: augmented reality—virtual reality— mixed reality——virtual city tour—guidance—tangibles asymmetric—collaborative virtual environment—

1 Introduction

Collaboration using mixed reality technologies gained great popularity in recent years. These technologies can help distributed team members to collaborate on a task without sharing the same physical environment [6]. In addition, these technologies can enhance the collaborative work environment with shared virtual content [16]. A virtual environment, in which multiple users are collaborating is also called "collaborative virtual environment" (CVE) [5]. Based on the user roles, a CVE can further be classified as an symmetric and asymmetric CVE, as defined by Feld & Weyers [7] and Ens et al. [6]. In a symmetric CVE, every user has the same role and thus has the same amount of authority over the environment as well as contribute a similar amount of information. Instead in an asymmetric CVE, one user role, called "primary" in this work, has a higher authority over the virtual environment and contributes most information. In contrast, a "secondary" user role only has minor authority and contributes a smaller amount of information. A typical asymmetric CVE is a telemaintenance system, where a remote expert (primary) is guiding a technician on site (secondary) through a maintenance task [11]. Here, the technician is in a more passive role due to the collaboration as she/he is only performing the task she/he is guided through by the primary expert.

*e-mail: feldn@uni-trier.de
†e-mail: weyers@uni-trier.de

In a literature study [7], Feld & Weyers identified a lack of research for systems, in which the primary is using augmented reality (AR) in asymmetric CVE. They mostly found examples of the secondary users using AR and the primaries using a common desktop environment [8] or a virtual reality (VR) system [12]. This leads to the question if there is a reason why and if the primary could also benefit from using AR in such a system. Therefore, the main research question is: "Can AR support a primary in a asymmetric CVE?"

While this work does not aim at answering this research question wholly, we present ongoing work including the development of a prototype, which should allow further research in this area. Therefore, we applied a user-centred development process to create this prototype of an asymmetric CVE for a virtual city tour. In this process, we formulated four hypotheses based on results of a first online survey and then further investigated these hypotheses with a (semi-)structured interview. In the system, the guide is taking the primary role to provide information to the visitors interactively. In contrast, the multiple visitors take the secondary role and are using virtual reality to visit a virtual city model of the ancient German city Trier [17]. The conducted empirical research indicates a certain benefit for city guides using AR to conduct virtual city tours, as AR and VR could improve the experience in such scenarios [13]. The presented prototype represents the second design iteration, where all prototypes have been qualitative evaluated in (semi-)structured interviews and online surveys as mentioned above.

The rest of the paper is structured as follows. Based on related work (section 2) and a preliminary study (section 3.1), we first define design requirements for the prototype in section 3.2. In consideration of these requirements, we then present our system design and the prototype implementation. For the evaluation of the prototype as part of the user-centered design process, we first conducted an online survey followed by a (semi-)structured interview with experienced city tour guides (see section 4). To conclude our work in progress, we discuss the results of the evaluation and possible future work in section 5.

2 Related Work

Using AR and VR in collaborative scenarios is a common topic in AR/VR research and, because of technology advances in the last years, it reached a state where current and future work can focus on nuances of collaboration rather than on technical difficulties [6]. The "Studierstube" system from Szalavari et al. [16] already proposed a symmetric CVE using AR to allow interactive scientific visualization in 3D for multiple users. AR was implemented with head-mounted displays (HMDs), which were tracked by an extra tracking device attached to the HMDs. A physical pen and panel were also tracked and augmented to allow an intuitive interaction with virtual objects. In a study, they concluded that this 3D exploration and manipulation of the 3D scientific visualization is superior to the typical 2D desktop-and-mouse interaction for working with these models collaboratively. A crucial part of this work is the AR-based augmentation of the pen and the panel, which led to an intuitive interaction with the system. Therefore, we designed our system using augmented physical objects as interaction props (i.e. tangibles [1]) to interact with the virtual environment.

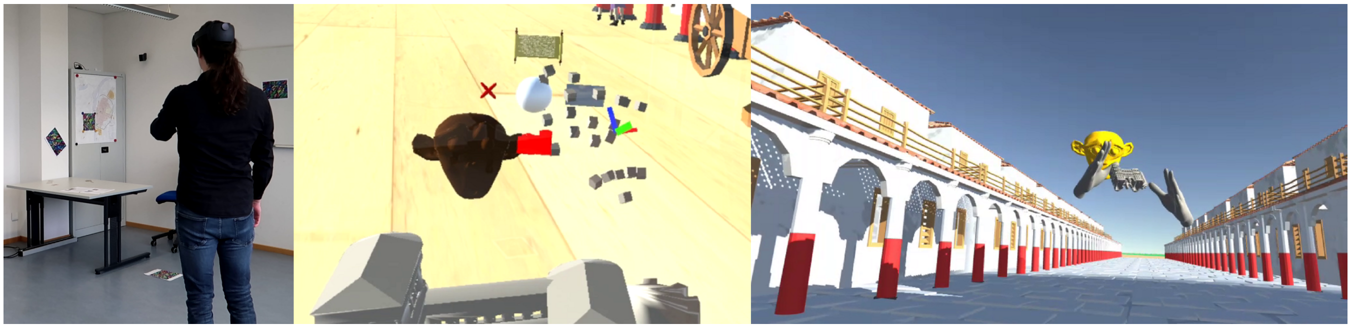


Figure 1: Overview of the prototype: The guide (left), her/his view (middle) and the view of a visitor (right).

One of the few asymmetric CVEs using AR for the primary is the "Dollhouse VR" system presented by Ibayashi et al. [10]. In their work, they implemented a prototype that allows the collaborative architectural design of a virtual room using AR and VR. The primary in this system is using an augmented table to view the virtual room from above in 2D (also called "tabletop" AR). The secondary is using VR to immerse herself/himself into the virtual room observable by the primary on the augmented table. A camera inside the augmented table allowed the creation of the illusion of a see-through ceiling inside the virtual room, which allowed the secondary to see the primary. Together, they were then able to communicate and redesign the room on their liking by interacting with the furniture. The primary was also able to interact with the furniture and re-position it using the touch-sensitive surface of the augmented table. Such manipulations were visualized to the secondary by a giant hand inside the virtual room. Their study showed that the system allows to redesign a virtual room intuitively and that it works as a communication medium between the users. However, they used a 2D tabletop AR implementation, which provides almost no immersion to the primary into the virtual room. This resulted in difficulties to perceive gestures and mimics of the other users, especially having eye contact or pointing in 3D. This inspired us to design and implement functionality in our system that enables the primary to engage with the secondaries on eye-level by allowing the primary to fully dive into the virtual city of the secondaries.

An example of a symmetric CVE in the setting of tourism respectively virtual museums is the "Lighthouse" system presented by Brown et al. [4]. This system allows multiple users to visit a museum simultaneously with three different technologies. The user can either visit the real museum with a PDA, a virtual copy of the museum in VR or visiting a web version of the museum using a browser. One requirement set by the authors was that all users must be able to talk to each other, share their position and orientation in the museum, and have a "common information space" as defined by Schmidt & Bannon [15]: a shared (virtual or physical) space that provides necessary information resources for the users to collaborate. Their study showed that the participants were able to interact collaboratively in the museum and enjoyed the social engagement. Furthermore, this work showed that a crucial part of visiting a museum is the presence of the other users collaborating with. They stated that while the users were aware of the other users' position and orientation, gestures were poorly supported by their system, but are important in a CVE. Therefore, we require our system to display other users in a way that supports gestures to make the communication between users easier.

In summary, we identified the following aspects in related work: **1.** It is beneficial to use real, augmented objects for an intuitive interaction with the virtual environment. **2.** It is vital for the primary to be able to immerse herself/himself into the virtual environment of the secondary. **3.** Visually representing the other users is essential for collaboration.

These points were thoroughly considered for the design of our prototype as described below.

3 System Design

3.1 Preliminary Study

In a preliminary study[1], a predecessor of our prototype was implemented and evaluated in a semi-structured interview. The AR part of the system for the primary was implemented with an ordinary smartphone, such that the primary constantly looked through the smartphone to see the augmented content. The primaries working space was designed to be her/his regular office desk. Via an augmented wooden shovel, she/he was able to view a section of the virtual city as a miniature on the desk, also called "mini-warpzone". The mini-warpzone was always centred around the current secondaries she/he was watching at. The focused secondary could be changed using an augmented physical paper map. The map showed the current secondaries present in the virtual city. They were rendered at their actual position in the virtual city. Executing a "pickup" metaphor with the wooden shovel, the primary was able to select a secondary. The selected secondary then got attached to the wooden shovel and the section of the virtual city around the secondary gets displayed in the mini-warpzone. When selected, the primary could show additional content to the secondary by using additional augmented wooden chips as tangibles. By selecting one of these wooden chips by clicking on the chip's representation on the smartphone touchscreen, she/he could select a 3D model to be displayed on the chip. Furthermore, the primary could interact with various objects. By clicking on them when shown in the smartphone screen, she/he could either show a text, play a video, or highlight the object to the secondary.

In the evaluation of this system as semi-structured interview, first, the participants had to perform a predefined tour in the primary's role followed by the interview with the experimenter. The group of participants consisted of students from an HCI course and colleagues from the authors' research group. The secondary was played by the experimenter. After finishing the task, the participants answered a set of questions about the system's usability in the form of an expert interview. The results of this interview suggest the following assertions.

All participants saw potential in the system and its design, but stated that the current implementation was too flawed to enhance the workflow in comparison to a typical desktop environment without the usage of AR. Regarding the usage of AR, they saw potential, as they could "better assess the behaviour of the user, because I [the participant] am more immersed", as one participant stated. The major downside of the prototype was the usage of the smartphone. Because the primary always had to hold the smartphone with one

[1] https://arxiv.org/abs/2101.02565

hand, she/he could only interact with the system with the other hand. "I see a big disadvantage in the mobile phone ..." and "... I think two [free] hands are clearly superior" were common statements. Further, the small screen of the smartphone provided a tiny interaction area to the primary so that "the buttons are close to each other", which made it difficult to interact with the 3D models. Some participants stated that they also had difficulties seeing where the secondary was looking at and suggested some visualized view cone.

3.2 Concept

The results of our preliminary study and the suggestions from related work led to a set of design requirements for our prototype. By addressing these design requirements, we aim at a prototype that should allow further research regarding the main research question. The design requirements are:

DR 1. Communication: **1.** The primary and secondaries need to communicate. **2.** The primary needs to see the secondary and her/his virtual environment. **3.** The primary needs to steer the attention of the secondary towards certain points of interest. **4.** The primary must be able to engage the secondary on eye level.

DR 2. Interaction: **1.** The primary needs to interact and manipulate the CVE in a intuitive and efficient way. **2.** The primary needs to show additional information to the secondary. **3.** The primary needs to use AR. **4.** The primary needs to potentially switch between multiple secondaries. **5.** Any context switches must be intuitive and non intrusive.

DR 3. Navigation: **1.** The secondary needs to freely move inside the CVE. **2.** The primary needs to interact with the secondary regardless their position in the environment. **3.** The primary must be able to guide the secondary to a location.

Based on these design requirements, a concept for the prototype was developed. One major aspect was again using AR due to **DR 2.3**. In addition, the prototype should use an HMD to project the augmented content such that we reach a high level of immersion. Furthermore, an HMD does not require any further handling once mounted, it is non-intrusive (compared to a smartphone), and allows the primary to use both hands to interact with the system. The secondary is using full immersive VR to engage in the system (creating a motivational learning experience [2]) and tracked controllers for navigation (meeting **DR 3.1**).

In the preliminary study, the only option for the primary to interact with the secondaries was via the pickup shovel. However, as we aim at making interaction and manipulation as intuitive and efficient as possible, a more direct and precise interaction with the secondaries was necessary. Therefore, we added a new virtual working space (VWS) the primary can switch to. In this VWS, the primary can enter the virtual city in real scale, which means that the primary gets fully immersed in the same virtual city as the secondary. Once entered, the primary can freely move in (via physical walking and steering using a bluetooth controller), as well as interact with the virtual city and the secondaries directly (**DR 1.4**). This newly added VWS will be further called "warpzone". Therefore, the primaries desk, now called "deskzone", is used to provide an overview over the whole virtual city, while the warpzone should allow the primary to engage with the secondaries directly on eye level. As stated above, the switches between these VWSs are a crucial part due to user experience and, thus, should be intuitive and easy to perform (meeting **DR 2.5**). Therefore, the deskzone and warpzone are physically separated. The deskzone is defined as an area around the desk, and the warpzone is everything around this area as shown in figure 2. Now, the primary only has to physically walk out of the deskzone into the warpzone and vice versa to switch between these two VWSs. The mini-warpzone from the preliminary study is still present on the primary's desk and bound to the same wooden shovel (**DR 1.2**). However it is no longer bound to a specific secondary and does not move with her/him anymore. Instead, the primary can change the

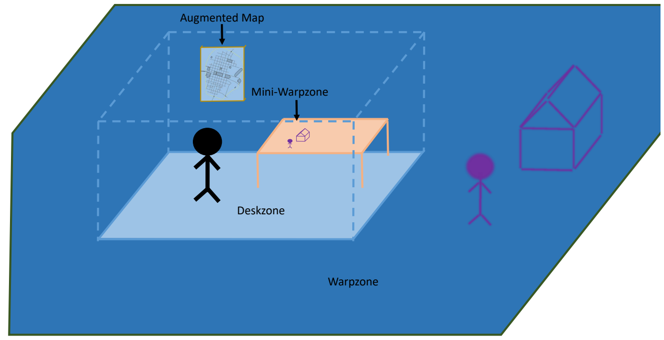


Figure 2: Draft of the physical separation between the deskzone and the warpzone

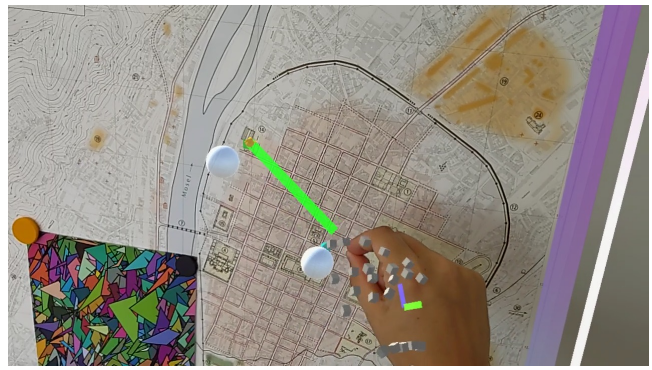


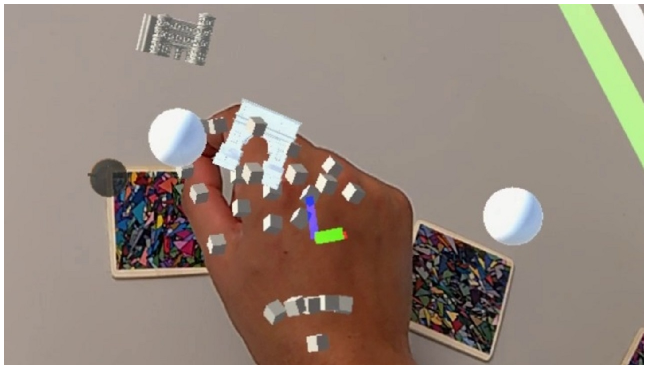
Figure 3: Guiding the secondary via the augmented map

section of the virtual city displayed by using the analog stick of the already mentioned bluetooth controller. Regardless the current VWS the primary is currently in, she/he can always communicate with the secondaries via a voice chat (**DR 1.1**).

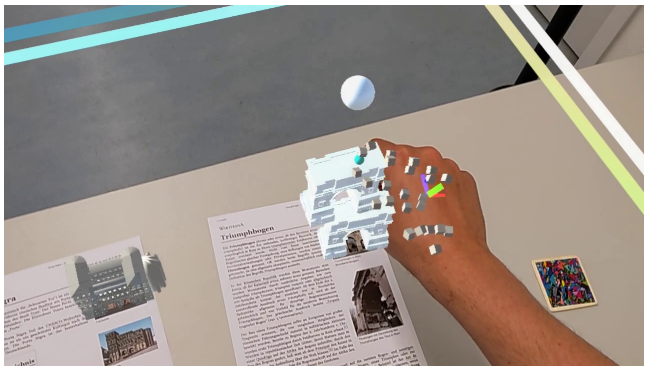
To display additional information for an object in the virtual city, the primary can interact with an object via a radial menu. Therefore, a small orb hovers above each object the primary can interact with. When the primary grabs this orb, various functions as 3D objects are displayed around the orb (see figure 1). When the primary moves the orb inside one of these objects after performing a grab gesture and subsequently releasing it, the corresponding operation is selected and executed. By the usage of hand tracking, we meet **DR 2.1**. The currently implemented functions in our prototype are: *displaying text*, *displaying a video* and *highlighting the object*, which meets **DR 1.3** and **DR 2.2**.

The augmented map as used in the preliminary study is also implemented in the current prototype. However, the primary cannot pick up a secondary, as the mini-warpzone is not bound solely to a secondary anymore. Each secondary and the mini-warpzone are displayed as a schematic representation on the map at their actual location in the virtual city. Above each representation, similar to the interactable objects, a small orb is hovering. By grabbing the orb of a secondary and releasing it above another location of the virtual city, a green arrow gets displayed in the view of the secondary to guide her/him to that selected location as seen in figure 3. When grabbing and releasing the orb of the mini-warpzones representation, the shown section of the virtual city by the mini-warpzone gets moved to this new location. Using this method, the primary can guide a secondary through the city (**DR 3.3**) and move the mini-warpzone to another location as needed, for example to another secondary (**DR 2.4** and **DR 3.2**).

To create additional objects in the virtual city, the primary has



(a) Selecting a model for an augmented chip



(b) Creating a new model with a augmented document

Figure 4: The selection and creation of new objects with the augmented wooden chips and augmented documents.

two options: Using augmented documents, which mimic normally used material of city guides, or small augmented wooden chips as tangibles. For the latter, the primary can bound objects (see figure 4a for all available objects) to them using the same orb-based radial menu as described above. A bound object then gets spatially registered to and displayed on top of the wooden chip. By grabbing this new object, the primary creates a copy of it, which she/he can then take into the warpzone to show it to the visitors, as shown in figure 4b. Each user can now grab this model to view it from any angle. The augmented documents are almost identical to the wooden chips, with the exception that they cannot be generically used. For example, one augmented document used in the study was a printed Wikipedia page of the Porta Nigra. Thus, the only model to be displayed is a model of the Porta Nigra on top of that document. By grabbing it, the primary can create a copy of it and use it as described above. Hence, the model selection via the radial menu is not present for the augmented documents.

The prototype was implemented with the Unity Engine 2019.4. As AR HMD for the primary, a Hololens 2 including hand tracking for all free hand interaction techniques was used. As VR headset, we used the Oculus Quest including controllers. The Networking between the users was implemented with the Photon Engine - Library 2.18 and the image tracking for the tangibles with the Vuforia Engine 9.6.3. The videos of the online-survey, which are describing the setup in more detail can be found here: `https://doi.org/10.5281/zenodo.4428762`. Please note that these videos are only available in German language.

4 Evaluation

For the evaluation of the prototype, we decided to investigate each element separately. In addition, the corona pandemic prevent us from conducting a laboratory study. To cope this problem, we first created an online survey. Based on its results, we formulated four hypothesis, which were further analysed due to relevance in a (semi-)structured interview with experienced city tour guides. This interview also revealed first promising insights regarding usefulness and relevance of the prototype for virtual city tours.

4.1 Online Survey

The online survey covered five sub questions related to the various elements (i.e. augmented map, tangibles, etc.) of the system design:

1: *Is the setup with the separation of warpzone and deskzone easy to understand?*
2: *Are the mini-warpzones easy to understand and easy to use?*
3: *Does the augmented map provide an overview of the virtual city and an efficient way to handle the secondaries and the mini-warpzones?*
4: *Does the warpzone provide an intuitive way to directly interact with the secondaries?*
5: *How does this virtual city tour compare to a real city tour?*

After watching a video showing and describing each of the prototype's elements, participants were asked to rate various usability-related items (e.g. complexity, difficulty and inconvenience) on a 5 point Likert scale for each sub question listed above.

The survey was sent to private and professional contacts. In addition, we were able to contact experienced city tour guides via the tourist information of the city Trier. 31 participants filled out the survey. The participants were 54.8% female, 45.2% male and 0% diverse or gave no information and were between 22 and 78 years old ($M = 44.0$ years, $SD = 16.5$). Most participant (93.5%) fully (61.3%) or rather agreed ($29,0\%$) that the videos helped them to understand our prototype. Also these results may imply that most participants had a good understanding of the prototype, we are confident that only from watching videos, participants neither fully grasp the prototype nor could reliably rate its usability. However, the survey was planned to formulate hypotheses regarding the usability of the prototype as an element of answering the overall research question.

The results indicate that the separation between the deskzone and the warpzone may increase the efficiency of the primary by providing an overview of the virtual city. In addition, it may offer the opportunity to directly approach a secondary and design the transition as intuitive as possible. Therefore, our first hypothesis reads:

(H1) *The physical separation of multiple different VWSs and an intuitive transition between these enhances the workflow of the primary to both interact directly with a secondary and get an overview over the whole virtual city.*

This also relates the answers we received regarding the mini-warpzones, which led us to the indication that the concept of the mini-warpzones does not need a redesign. In this context, the usage of tangibles for creating new objects was also well received. This led to our next hypothesis:

(H2) *The usage of tangibles does support the primary on interacting with the system, as well as interacting with a secondary.*

The augmented map was the most prominent physical object, which was augmented with virtual objects. Based on the results of the survey, we think that linking the functionalities of both the real map (getting an overview) and the virtual objects (interacting with a secondary or a warpzone) can improve the intuitiveness of the system, which leads to the hypothesis:

(H3) *Combining functionalities of real objects with functionalities of virtual objects via augmentation leads to a more intuitive and efficient way for the primary to guide a secondary and get an overview over the virtual city.*

Regarding the differences between the tangibles and the augmented documents, there was no clear indication that the separation between these two is necessary. However, considering the context of the online survey, we see need to evaluate this aspect further.

The evaluation of the overall system indicates that while the participants see potential, some argued that it is not an alternative for typical city tours in the near future. This leads to the hypothesis:

(H4) *A realistic presentation of the virtual city has an impact on the acceptance of the virtual city tour.*

The collected data and, thus, the hypotheses set up above have a few limitations. As already mentioned, the participants were only able to watch prerecorded videos of the prototype and, thus, could not try the prototype hands-on. Further, the data was collected in an uncontrolled environment.

4.2 (Semi-)structured Expert Interview

To address these issues, we invited 10 experienced city tour guides (50% male, 50% female) and conducted an expert interview to further evaluate the hypotheses and to gain further feedback regarding the quality of the prototype. The interview combined structured (using a 5 point Likert scale-like questionnaire) and semi-structured (pre-defined questions with open answers) elements. Each invited city tour guide received an introduction to the prototype by the experimenter and got time to get familiarized with the prototype hands on. After the introduction and some free time, a short virtual city tour was reenacted. By performing a set of tasks, all elements of the prototype are used again by the participant following a predefined schema. Two examples of tasks are "Convey information about the old Porta Nigra to the visitor. Use media content to do this." or "Guide the visitor to the Horrea.". These tasks were explicitly formulated without the keywords like "augmented map" or "radial menu" to not influence the participant which tools should be used to solve the task. During the introduction, familiarization and task execution phase, the experimenter played the role of the secondary. To regard physical and mental fatigue, we decided for a maximum time limit of two hours for each participant in total. The whole interview was conducted in German language.

After finishing all tasks, we asked the participants (in the structured part of the interview) various questions about their experience with the prototype and their opinion in regards to the hypotheses we formulated. In case the participant agreed, the whole interview was audio recorded such that every comment and statement was captured and could be transcribed. Hypotheses H1, H2, and H3 combine various sub-hypotheses and had to be separated to be better addressable in the interview. Each of these sub-hypotheses were addressed by one structured question in the interview, which was answered by means of five items, answered on a 5 point Likert-scale each. The items were inspired by the system usability scale (SUS) [3]. As an example, figure 5 shows the results addressing H1, which was split into four different questions. This procedure was repeated for H1 to H4, which resulted in nine questions (H1: four, H2: two, H3: two, H4: one). In addition to the nine questions, we asked for possible alternatives to our concept of two VWSs and regarding the metaphor to switch between them, the usability about the radial menu, and missing interactions with the objects and the visitors.

3 of the 10 participants were left out of the results, as either to many technical problems occurred or the time limit did not allow for the reenacted virtual city tour. However, their comments and statements were still taken into account excluding the feedback on the nine hypotheses related structured questions. The questions and the comments from the participants were translated into English by the authors. Further, the results of each negative item *unnecessarily complex*, *inconsistent* and *cumbersome* were inverted in each figure to match the color connotation. These items were marked with a *(not)*. The complete online survey and the (semi-)structured interview together with the results can be found here: `https://doi.org/10.5281/zenodo.4455393`.

Figure 5 shows the answers of the four usability questions about

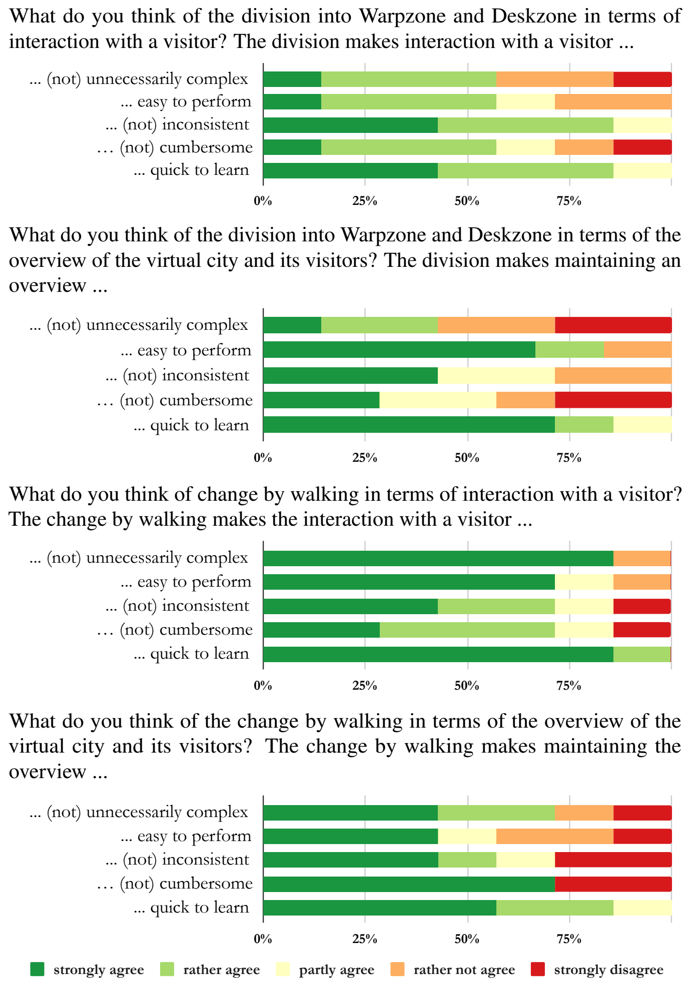


Figure 5: Usability of the separation into two VWSs. (H1)

H1 regarding the separation into two VWSs and the transition via walking. Especially the results for the "easy to perform" and "quick to learn" items indicate a tendency to high usability for the workflow with two VWSs. Some participants stated that they had difficulties to orient themselves after the switch into the warpzone and that the workflow is a bit complex and cumbersome to use. While the results for the inconsistency indicates a shift to a high usability, we do not interpret these results, as many participants stated they had difficulties to understand what inconsistency means, even after an explanation by the experimenter. Therefore, we see potential in this hypothesis to guide future research in terms of the general research question.

The results of items addressing H2 are similar to the results of H1, as seen in figure 6. The responses for "easy to perform" and "quick to learn" are promising regarding a high level of usability. However, a few participants did strongly, rather or partly agree that the workflow using tangibles are unnecessary complex and cumbersome. We speculate that the agreement of the "cumbersome"-statement is a result of frequently lost tracking of the tangibles due to technical reasons or occluded markers by participants. However, the idea of using physical props to interact with the system was well received, which was also supported by the participants' statements during the introduction and the reenacted city tour. This indicates that, after enhancing the tracking of the tangibles, H2 could still help us answering our main research question, and therefore we want to further investigate H2 in the future.

When we asked whether replacing the real city map with a virtual

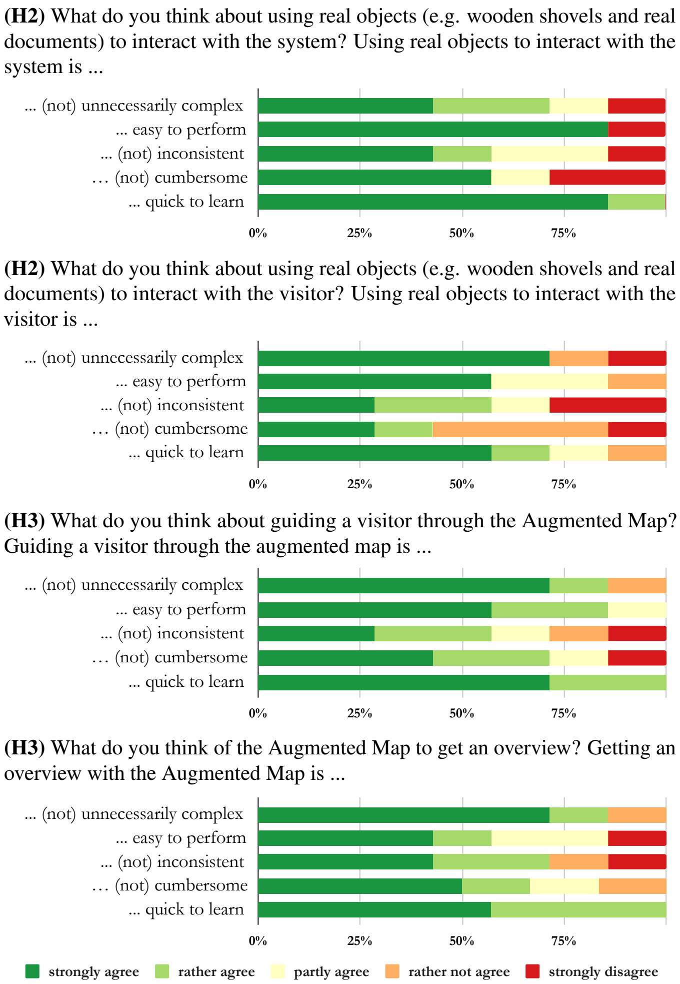


Figure 6: Usability of the tangibles (H2) and the augm. map (H3).

map would have any positive or negative impact on the workflow, 57.1% stated it would have no impact and 28.5% that it would make the workflow easier. Only 14.3% think the workflow would be more difficult. This indicates that a rejection of H3 could be possible, as the combination of the real map with the virtual objects was not well received. However, the function of the real map was just gaining an overview and, thus, had no active role. Therefore, this hypothesis could be evaluated with a more active role of the map, which could be gained by including additional functionality for the passive objects in the scenario. The results of the usability questions, as seen in figure 6 indicate a high usability of the virtual functionalities of our prototype. Therefore, we think that the role in the virtual environment of the map is useful, however the role in the real environment is to passive to either accept or reject H3.

In the online survey, many participants stated that a realistic representation of the virtual city is essential to an acceptance of the city tour, as the visual representation seem to have an impact on the presence in a VE [9]. However, results gained in the (semi-)structured interview were not so clear. With two participants each stating "rather yes", "part-part" and "rather no" and only one stating "yes very", no clear indication for a potential rejection or acceptance of H4 can be expected. In addition, many participants suggested the alternative setting of a merged virtual city tour: a real city tour gets enhanced by virtual objects. Thus, the visual representation of the virtual city is not as important as in our setting and these participants stated that they answered this question in regards to this merged setting and not in terms of our virtual city tour setting. Furthermore, the representation of other users was indicated to be too simple in our prototype. The participants referred to missing facial expression and gestures from both, the guide and the visitors, which is necessary for a successful communication between the users.

Regarding the separation of the two VWSs, all participants stated that they think the separation into two VWSs is useful. However, six participants either suggest to merge the two VWSs into the same physical space or change the VWS switch to a less physical metaphor. They argue that this would make it less tedious to perform a complete city tour in its full duration. However, other participants stated that they would like to keep these physical movements, as it would keep the primary more engaged during the virtual city tour.

When asked about possible alternatives to the two VWSs, two participants suggested to attach the deskzone to the primaries body like a hawker's tray, which would make a switch between the VWSs unnecessary. However, the concept of tangibles has to be redesigned when no physical desk is present anymore. An augmented real hawker's tray could be a solution, similar to the work by Pham & Stuerzlinger [14].

5 Discussion & Future Work

The prototype is work in progress. Therefore, we can not make any final statements about the initial research question. Further, all statements about the hypotheses are not backed up with a quantitative study, which is the main downside we want to address in future research. Further the measuring scales could be adjusted to lead to more conclusive results. However, we conclude that the online survey followed by the (semi-)structured interview helped us to get a picture about the various aspects of such a system and the needs of a potential virtual city tour guide and lead to potential future research guided by the investigated hypotheses.

For H1, the results and the comments of the experts indicate to further investigate the usage of multiple VWSs, specifically the aspects of separation and transition, as we think that H1 could be accepted by further quantitative empirical research. We suggest to evaluate multiple types of VWSs and various separation and transition metaphors. As an example the suggested "hawker's tray" could be a valid alternative to our approach of separating the VWSs physically.

After the (semi-)structured interview, we still think that further research on using tangibles could lead to an improved experience for the primary and offer insight due to H2. However, many technical issues have to be solved with a focus on the identified tracking issues. One possible solution could be to use three dimensional objects rather than just flat wooden chips. This could not only improve the usability, as they are easier to grab, but also improve the tracking as they can be tracked from multiple angles.

The results for H3 indicate that future quantitative studies may lead to rejection of this hypothesis. However, the augmented physical map had no active role, which may have cause this feedback. Therefore, we suggest to change the hypothesis to: *Combining **active roles** of real objects with **active roles** of virtual objects via augmentation leads to a more intuitive and efficient way for the primary to guide a secondary and get an overview over the virtual city.*

Finally, H4 seems to be adapted as well, as the results indicate that an unrealistic representation of the virtual city may only have a negative impact in case of a full immersive virtual city tour. Many participants suggested the setting of a mixed city tour and stated that a realistic representation in such a system would be not as important. This leads to our suggestion to reformulate H4 into: *A realistic presentation of the virtual city has an impact on the acceptance of the virtual city tour **in a fully immersive virtual city tour**.*

We think that the proposed setting of the mixed virtual city tour is an interesting idea but differs significantly from our work. To prevent a full redesign, we first want to keep the setting as it is and evaluate the hypotheses with the prototype in future work.

Acknowledgments

The authors wish to thank Ursula Bartmann and Caroline Baranowski from the tourist information Trier for their support and all participants of the online survey and interview.

References

[1] M. Billinghurst, H. Kato, and I. Poupyrev. Tangible augmented reality. *ACM SIGGRAPH ASIA 2008 Courses*, 01 2008. doi: 10.1145/1508044.1508051

[2] C. Borst, N. Lipari, and J. Woodworth. Teacher-guided educational vr: Assessment of live and prerecorded teachers guiding virtual field trips. pp. 467–474, 03 2018. doi: 10.1109/VR.2018.8448286

[3] J. Brooke. Sus: A quick and dirty usability scale. *Usability Eval. Ind.*, 189, 11 1995.

[4] B. Brown, I. Maccoll, M. Chalmers, A. Galani, C. Randell, and A. Steed. Lessons from the lighthouse: Collaboration in a shared mixed reality system. 10 2002. doi: 10.1145/642611.642711

[5] E. F. Churchill et al. Collaborative virtual environments: Digital places and spaces for interaction. *Computer Supported Cooperative Work*, 2001.

[6] B. Ens, J. Lanir, A. Tang, S. Bateman, G. Lee, T. Piumsomboon, and M. Billinghurst. Revisiting collaboration through mixed reality: The evolution of groupware. *International Journal of Human-Computer Studies*, 131:81 – 98, 2019. 50 years of the International Journal of Human-Computer Studies. Reflections on the past, present and future of human-centred technologies. doi: 10.1016/j.ijhcs.2019.05.011

[7] N. Feld and B. Weyers. Overview of collaborative virtual environments using augmented reality. In *Mensch und Computer 2019 - Workshopband*. Gesellschaft für Informatik e.V., Bonn, 2019. doi: 10.18420/muc2019-ws-617

[8] K. Gupta, G. Lee, and M. Billinghurst. Do you see what i see? the effect of gaze tracking. *IEEE Transactions on Visualization and Computer Graphics*, 22:1–1, 07 2016. doi: 10.1109/TVCG.2016.2593778

[9] J. S. Hvass, O. Larsen, K. B. Vendelbo, N. C. Nilsson, R. Nordahl, and S. Serafin. Visual realism and presence in a virtual reality game. *2017 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON)*, pp. 1–4, 2017.

[10] H. Ibayashi, Y. Sugiura, D. Sakamoto, N. Miyata, M. Tada, T. Okuma, T. Kurata, M. Mochimaru, and T. Igarashi. Dollhouse vr: A multi-view, multi-user collaborative design workspace with vr technology. In *SIGGRAPH Asia 2015 Emerging Technologies*, SA '15. Association for Computing Machinery, New York, NY, USA, 2015. doi: 10.1145/2818466.2818480

[11] F. Leutert and K. Schilling. Projector-based augmented reality for telemaintenance support. *IFAC-PapersOnLine*, 51(11):502 – 507, 2018. 16th IFAC Symposium on Information Control Problems in Manufacturing INCOM 2018. doi: 10.1016/j.ifacol.2018.08.368

[12] O. Oda, C. Elvezio, M. Sukan, S. Feiner, and B. Tversky. Virtual replicas for remote assistance in virtual and augmented reality. pp. 405–415, 11 2015. doi: 10.1145/2807442.2807497

[13] S. Park and B. Stangl. Augmented reality experiences and sensation seeking. *Tourism Management*, 77:104023, 2020. doi: 10.1016/j.tourman.2019.104023

[14] D.-M. Pham and W. Stuerzlinger. Hawkey: Efficient and versatile text entry for virtual reality. In *25th ACM Symposium on Virtual Reality Software and Technology*, VRST '19. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3359996.3364265

[15] K. Schmidt and L. Bannon. Taking cscw seriously: Supporting articulation work. *Computer Supported Cooperative Work*, 1:7–40, 03 1992.

[16] Z. Szalavári, D. Schmalstieg, A. Fuhrmann, and M. Gervautz. "studierstube": An environment for collaboration in augmented reality. *Virtual Real.*, 3(1):37–48, Mar. 1998. doi: 10.1007/BF01409796

[17] D. Zielasko, N. Feld, C. Flemming, P. Lungershausen, A. Morgenthal, S. D. Schmitz, T. Mattern, and B. Weyers. Towards preservation and availability of heterogeneous cultural heritage research data via a virtual museum. In B. Weyers, C. Lürig, and D. Zielasko, eds., *GI VR / AR Workshop*. Gesellschaft für Informatik e.V., 2020. doi: 10.18420/vrar2020_6