



Article

There is Always a Way: Organizing VR User Tests with Remote and Hybrid Setups during a Pandemic—Learnings from Five Case Studies

Sanni Siltanen ^{1,*}, Hanna Heinonen ¹, Alisa Burova ², Paulina Becerril Palma ^{1,2}, Phong Truong ¹, Viveka Opas ¹ and Markku Turunen ²

¹ KONE Corporation, Myllykatu 3, 05830 Hyvinkää, Finland; hanna.heinonen@kone.com (H.H.); paulina.becerrilpalma@tuni.fi (P.B.P.); phong.truong@aalto.fi (P.T.); viveka.opas@kone.com (V.O.)

² Faculty of Information Technology and Communication Sciences, Tampere University, Kalevantie 4, 33014 Tampere, Finland; alissa.burova@tuni.fi (A.B.); Markku.Turunen@tuni.fi (M.T.)

* Correspondence: sanni.siltanen@kone.com

Abstract: (1) COVID-19 pandemic restrictions caused a dramatic shift in research activities, forcing the adoption of remote practices and methods. Despite the known benefits of remote testing, there is limited knowledge on how to prepare and conduct such studies in the industrial context where the target users are experts and company employees. (2) In this article, we detail how we organized VR user tests in five industrial cases during the pandemic, focusing on practicalities and procedures. We cover both on-site testing, including disinfecting and other safety protocols, as well as remote and hybrid setups where both remote and on-site participants were involved. Subject matter experts from eight countries were involved in a total of 22 tests. (3) We share insights for VR user test arrangements relevant to the pandemic, remote and hybrid setups, and an industrial context, among others. (4) Our work confirms that with careful planning it is possible to organize user tests remotely. There are also some limitations in remote user testing, such as reduced visibility and interaction with participants. Most importantly, we list practical recommendations for organizing hybrid user tests with safety and disinfecting procedures for on-site VR use.

Keywords: user test; remote user test; hybrid user test; remote facilitation; VR; virtual reality; remote collaboration; disinfecting; remote observation; remote interviews; collaborative VR; user test planning; user test organization



Citation: Siltanen, S.; Heinonen, H.; Burova, A.; Palma, P.B.; Truong, P.; Opas, V.; Turunen, M. There is Always a Way: Organizing VR User Tests with Remote and Hybrid Setups during a Pandemic—Learnings from Five Case Studies. *Multimodal Technol. Interact.* **2021**, *5*, 62. <https://doi.org/10.3390/mti5100062>

Academic Editors: Fotis Liarokapis and Cristina Por-talés Ricart

Received: 19 August 2021

Accepted: 24 September 2021

Published: 11 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The COVID-19 pandemic has caused critical economic and social disruption all over the world [1,2], leading to a significant transformation of work and research-related activities [3,4]. While the regulations to follow “work from home” and, similarly, “research from home” [5] come from the governmental level, academia and industry research organizations have faced several challenges in managing their traditional workflows. Restrictions on physical contact and social isolation have massively affected the research projects related to the use of Extended Reality (XR) and immersive environments, where in-lab experiments with the presence of a moderator and a participant in one space accounted for more than 70% of the overall user study methods [6,7]. Similarly, other research activities, for instance based on the usage of mobile devices, have faced similar complications, and require additional efforts to be conducted.

XR technologies, which include augmented, mixed, and virtual realities (AR, MR, and VR), is a hot topic of HCI (human-computer interaction) research among academic and industrial practitioners. XR technologies offer unlimited benefits in the industrial context via blurring the boundaries between what is real and what is virtual, thus extending the possibilities of traditional technology appliances and supporting industrial

tasks. The cumulative evidence has proven the benefits of XR for industrial needs; XR has shown increased performance in facilitating collaboration and planning in construction projects [8–11] and immersive training of novice maintenance technicians [12], as well as the support of the whole product lifecycle in the manufacturing field [13–15]. Furthermore, the flexibility of XR to facilitate remote coordination and communication is even more pronounced during the pandemic [16,17].

Due to the above-mentioned advantages, large international corporations, such as KONE [18], perform their own research activities to investigate the appliance of novel technologies to their field. Such studies are usually conducted either internally or in collaboration with academia, targeting the company's novice and expert employees from different counties. The lockdown and travel restrictions caused a notable disruption in industrial and academic research activities, especially with topics that involve practical experiments including XR or mobile devices [19–21]. To continue their work activities, industrial researchers were forced to rapidly shift towards adopting remote practices and methods.

Remote usability testing and user studies are not novel concepts for the research world [22,23]. The effects of remote setups and related challenges has been investigated in academia since the early 2000s, due to potential benefits in cost and resource reduction [24]. The existing studies on remote testing demonstrated no significant variance between physically facilitated tests and remote tests; however, several challenges, such as research validity and lack of control, were also highlighted [25]. In XR research, remote studies have also been experimented with, targeting the growing group of HMD (head-mounted display) owners [26,27]. Nevertheless, there is neither clear guidance on how to conduct remote studies, nor an extensive body of research to base decision-making processes on when designing for a remote experiment. Furthermore, there is limited knowledge of conducting remote user testing in the industrial context, where the target user group is limited to experts and company employees. To fill this gap and demonstrate the practicalities of conducting remote user studies in the industrial context, our article details five industrial use cases, performed by a large international corporation in the field of manufacturing and maintenance during the pandemic restrictions. Our hypothesis is that it is possible to conduct user studies with remote and hybrid setups. The hypothesis is confirmed with the five use cases of user studies, which were successfully carried out during the pandemic. Two out of five studies were performed in collaboration with academia, based on the framework of remote academia-industry collaboration [28]. All five use cases provide a unique perspective to the practicalities of remote user studies, detailing how to approach the preparation phase and the remote processes of handling the experiments and safety aspects related to COVID-19 restrictions. The article, therefore, answers the following research question:

RQ: How can VR user studies be conducted with remote and hybrid setups in the industrial context?

The results of the remote case studies are presented in three categories to support the future implementation of remote experiments by industrial and academic practitioners: (1) insights that are especially important because of the pandemic, (2) insights that are especially important because of remote and hybrid setups, and (3) insights that are especially important because of the industrial context. In addition, we share (4) insights regarding conducting remote and hybrid interviews. The results are further discussed from the perspective of (5) general insights on user testing, and things related to (6) testing new, fascinating technology. Furthermore, we describe (7) hygiene and safety protocols for conducting VR user tests during a pandemic.

2. Related Work

The COVID-19 pandemic isolation and other restrictions have massively affected research activities, which has led to the adoption of remote methods and practices. However, remote research was already being investigated long before the pandemic situation; therefore, in this section we briefly describe the overview of remote studies, its drawbacks and benefits, and practical cases of using remote methods after the pandemic restrictions came into effect.

2.1. Brief History and Benefits of Remote Experiments

The remote experiment is not a novel concept; the idea of conducting remote usability testing started to develop over two decades ago [22,23,29]. Back in 1996, several types of remote evaluation were described by Hartson et al. [22], demonstrating the feasibility and validity of desktop video conferencing and remote evaluations.

Due to the promising advantages of remote testing—costs and time savings as well as freedom of facility [24], which promised to diminish geographical and organisational barriers [30]—various studies contributed to the comparison of remote experiment practices versus traditional lab settings and tried to explore their influence on the validity and procedure protocol, as well as other possible drawbacks [24]. For instance, the study by Andreasen et al. [30] compared three methods for remote usability testing to a conventional laboratory test and indicated the potential of the synchronous remote method. In comparison with the in-lab testing procedure, the synchronous remote method (moderated via remote desktop connection and a video conferencing test) did not show much variance in the number of identified mistakes and test duration. Nevertheless, asynchronous methods of remote studies were found to be less efficient and more time-consuming. Furthermore, the study by Madathil and Greenstein [31] compared a traditional lab setup to the two methods of synchronous remote usability testing: via 2D video conferencing, namely WebEx, and via a 3D virtual world. Their results demonstrated, on one hand, similar effectiveness of the test in the 3D virtual world in terms of test duration and number of higher severity errors, coupled with a slightly larger identified number of lower severity errors. On the other hand, despite a higher level of involvement during the test in the 3D virtual world, their results showed a higher level of workload in both remote conditions and identified delays as the major disadvantage of such a study. A more recent study [32] investigated the novel method of utilizing a collaborative virtual environment for usability testing and demonstrated the usefulness of the method due to higher user involvement and immersion. They also pointed out the difficulty of instructing users when their point-of-view was not visible to moderators.

In the last couple of years, due to the growth of market-available HMDs, the body of remote research on XR technologies has started to develop as well. For instance, the paper by Radiah et al. explored the challenges and opportunities of running remote VR studies, defining these as “studies taking place physically in users’ homes but not being restricted with regard to the virtual setting” [26] (p. 4). They further discussed three approaches of remote VR studies—standalone, app store, and VR platform—and concluded with a framework for running remote VR studies.

The effects of remote data collection methods (such as face-to-face versus technology-mediated interviews) were also studied thoroughly [33,34]. The studies reported that despite the lack of social presence [33], technology-mediated interviews may be more comfortable to the interviewees due to familiar surroundings and may reduce reactivity bias [33,35]. Furthermore, the longitudinal study by Weller [36] concluded that remote interviews with a video connection and physical separation may facilitate emotional connection via the increased sense of ease and limited pressure.

Among other challenges of remote studies, the following were mentioned throughout the research articles: lack of control and non-verbal communication, unpredicted affecting variables, cultural differences [37], technical issues, and network limitations. To address these and further discover remote methods, multiple papers proposed a variability of

tools and methods to maintain remote testing procedure [38–40] and demonstrated how to maximize the number of participants via crowdsourcing [41,42] or via making the research prototypes public on app stores [43]. Despite the potential of remote research methods, the research practitioners still preferred traditional experiments and moderated in-lab user studies, especially in the research projects with XR technologies [7].

Industrial user studies should include several experts and professionals with a domain knowledge to gain valid research output [28]. In normal conditions, people are brought together for user testing. However, this possibility completely vanished due to the pandemic isolation, which caused all research communities to adopt remote methodologies.

In the industrial context, a diverse group of industrial employees is often used in remote experiments to assist and facilitate sessions. Their expertise is primarily in subject matters and not in research methods, which brings challenges to organizing the remote sessions. Safety issues are important when carrying out experiments in the industrial environment and pose some limitations in the test setup. Often, a remote assistant is also needed to take care of safety [28].

2.2. Mobile and XR Remote Experiments during COVID

With the spread of COVID-19 related restrictions, many authors proposed methods to maintain remote XR research [44–47] and demonstrated the methods and practicalities of remote user studies, which we present in this sub-section.

Most authors followed the above-mentioned framework by Rivu et al. [26], and further suggested overcoming pandemic restrictions via targeting and attracting HMD-owners via social VR systems. On the other hand, Steed et al. [27] highlighted the challenges of recruiting participants owning HMDs: even though this potentially increases the participant pool, being early adopters, they may be biased toward the potential of VR technologies and thus, may affect the ecological validity negatively. Among other proposed methods, Steed et al. [44] mentioned a short-term solution to recruit personnel and maintain collaboration between VR labs as well as a long-term solution to purchase and lend equipment to the users. A similar idea was presented by Estrada and Prasolova-Førland in their article on running XR research activities during COVID-19 in Norway [45]. These approaches might solve the isolation problems in academic experiments; however, they do not fully address the needs of industrial research.

Other papers shared their remote practices and demonstrated the scope of unpredicted challenges that are essential to acknowledge prior to planning for remote research. For instance, an article by De Ponti et al. [48] reported the procedures of online learning with access to a VR training platform among medical students, which resulted in one-third of the students facing technical and user interface issues during the learning process. In the article by Albastaki et al. [49], the use of virtual experience prototypes for robotics evaluation was tested with expert users via online Zoom sessions. The results demonstrated that the think-aloud protocol was adopted with ease during the remote procedure, whereas screen-sharing and video recording were sufficient to observe the participants. On the other hand, good internet connection and a prior knowledge of the conferencing tool is critical to avoid a prolonged session time or misleading results.

To conclude, despite the potential of remote research practices, they were not widely adopted [7] before a recent shift towards research and work from home due to COVID-19 pandemic restrictions. In the case of academic research, various practices and tools were presented prior to the pandemic to address the challenges of remote research, which in turn could involve a wider audience and positively affect research outcomes. In the case of industrial research, where a group of possible participants is limited to experts and people with field experience, the mentioned practices are useless. In order to continue industrial research activities, practitioners were forced to overcome social restrictions in other ways and reach their employees at homes. To address the challenges of remote industrial practices of research, we share five case studies and detail how the practicalities were handled. User test setups for each case are described and insights from the cases

are gathered. This article presents methods for addressing the pandemic restrictions in an appropriate way, e.g., following the proper safety regulations and disinfection procedures and running hybrid-facilitated studies and interviews. Due to the industrial context, in remote and hybrid setups, company employees with no research background are used in assisting user tests. Challenges arising from this setup are discussed as well.

3. Case Descriptions

We organized remote user sessions in five different industrial use cases during the 2020–2021 pandemic. All cases concerned collaboration in VR. The total number of organized user test sessions was 22, and several pre-sessions were organized in addition to this. All together we involved subject matter experts from eight countries: Finland, India, China, Germany, Indonesia, Malaysia, the United States, and the United Arab Emirates. Some participants attended user tests in more than one use case, and, therefore, the total number of persons involved as participants is somewhat lower than sum of all participants. Table 1 gives a summary of all cases. Case 1 and Case 2 were KONE internal studies, and Case 3 and Case 5 were part of Human Optimized XR (HUMOR) [50] project activities in collaboration with Tampere University and KONE. Case 4 was part of the KEKO Smart Building Ecosystem project [51] activities at KONE.

Table 1. Summary of user test cases.

Case	Description	Number of Participants	Number of User Test Sessions	Type of Participants: Remote/Onsite HMD, Desktop, Teams, Smartphone
Case 1	VR collaboration for maintainability	3	6	Remote HMD, Desktop
Case 2	VR collaboration in general	6	1	On-site HMD, remote HMD, Desktop
Case 3	Asynchronous VR collaboration for maintenance method development and documentation	7	7	On-site HMD, remote HMD
Case 4	VR collaboration for complex machine room planning	9	3	On-site HMD, Remote HMD
Case 5	VR collaboration for maintenance method development and documentation	20	5	On-site HMD, Remote HMD, Teams

The terminology used in the literature for participants and researchers organizing user testing varies. For clarity we have summarized the terminology used in this paper in Table 2.

Table 2. Description of terms used in this paper.

Term	Definition
On-site facilitator	Facilitator who is in the same room with a participant.
Remote facilitator	Facilitator who is in a different physical location than the participant, can wear an HMD if facilitating inside VR, or facilitates in the real world.
On-site assistant	Assistant who is in a remote location with the participant, helping with the equipment and practical arrangements. Often a company employee without a research background.
Chat facilitator	Facilitator who facilitates the Teams discussion with Teams participants and assigns tasks.
HMD participant	Participant who wears a VR headset and hand controllers.
HMD _e participant:	VR participant whose view is shared with the Teams participants, e = eye.
Teams participant	Participant who participates via Teams.
Desktop participant	Participant who joins the VR session with the desktop UI.
Remote HMD participant	VR participant who is in a remote location.
Remote participant	Participant who is in a remote location
Remote observer	Researcher who observes remotely via Teams
Interpreter	Person who translates speech orally to participants and facilitator.
Pre-test	A “dress rehearsal” session, where the whole setup and flow of the user test session is carried through in preparation for real user tests.

3.1. Case 1: Collaborative VR for Maintainability Review

Case description: In this project, we investigated the benefits of multi-user virtual reality for a collaborative maintainability review during an early product development phase [52]. We used two commercially available VR environments: DesignSpace [53] and Glue [54]. Both environments also had support for desktop participants. The VR devices used were the Samsung Odyssey (facilitator), the HP Reverb (participant), and the HTC Vive (participant, observers)—see Figure 1.



Figure 1. A participant attending a test session at a remote assistant's home in Case 1.

In this project, a Lean start-up methodology was used [55]. Collaboration methods, tasks, and VR environments were developed in iterative sprints. In total, six user test sessions were conducted. Three sessions were held in Glue and another three in DesignSpace (see Figure 2).



Figure 2. DesignSpace screenshots from user tests in Case 1.

Pandemic restrictions: This project started in February 2020, and the generic introduction of VR environments for some of the users was done face-to-face in the office premises before the restrictions started. In March 2020, the pandemic situation worsened in the Uusimaa area, Finland, where this research took place, and restrictions were in place to prohibit entering the office premises. The company allowed a limited number of co-workers to meet in person if needed. Thus, all user tests were conducted remotely from people's homes (as shown on Figure 1). The project ended in July 2020.

Test setup: All sessions started with a Teams call, where the participants were introduced to the VR task, and an on-site assistant instructed how to start the VR environment. The facilitator explained how the hand controllers, menus, etc. work, and showed videos and screenshots from VR during the starting session. The Teams session was also used when practical help was needed with the VR devices or the VR environment. In tests with Glue, there was a remote facilitator inside the VR wearing an HMD, 1–2 HMD participants, and 1–2 desktop participants depending on the session; the total number of participants was three in each session. In tests with DesignSpace, there was a remote facilitator inside the VR wearing an HMD and 2 HMD participants, 1 desktop participant, and 1–2 observers (HMD). The same three participants attended all iterative sessions (Figure 3).

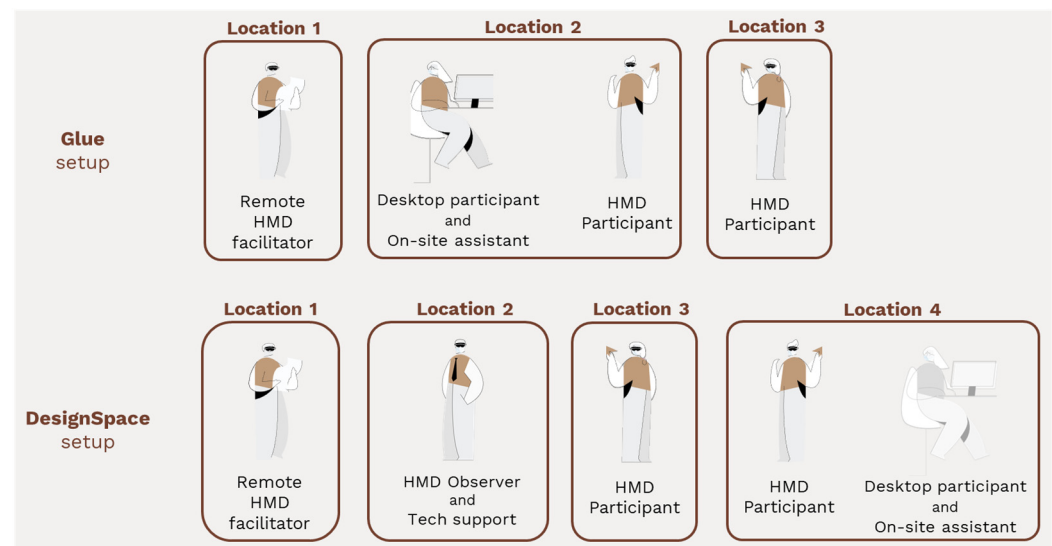


Figure 3. Different test setups used in Case 1. In DesignSpace sessions, one HMD participant joined alone from their location. The other HMD participant joined either alone or from the same location as the desktop participant, who acted as an on-site assistant.

Case-specific challenges: In order to include necessary subject matter experts despite the pandemic restrictions, one of the HMD participants was sent a VR set to their home and was instructed remotely via Teams on its use. The participant had no previous experience of setting up a VR environment (e.g., calibrating the play area and updating the VR software). The participant was a 60+-year-old expert on industrial maintenance and maintenance methods development, with hands on experience in the field.

3.2. Case 2: Multi-User VR in Global Department-to-Department Collaboration

Case description: One separate user test session was conducted to evaluate multi-user VR collaboration in a global department-to-department setting. In this session, the Glue environment [54] was used (see Figure 4).

The session focused on identifying the possibilities of VR for global collaboration and potential use situations and use cases. The participants had different professional backgrounds: maintenance methods development, technical documentation, and learning and development. Participants were from two different countries with different levels of previous VR experience. This work was KONE's internal development activity related

to the Human Optimized XR (HUMOR) project. The results of this user test have not been published elsewhere, but the results were taken into account in the planning of the following cases.

Pandemic restrictions: This test was carried out in July 2020. The company restrictions restricted access to the office premises, but a limited number of co-workers were allowed to meet in person if needed.



Figure 4. Snapshot of a laptop running Glue in desktop mode during a user test in Case 2.

Test setup: In this session, we had a remote HMD facilitator inside the VR, three HMD participants, and three desktop participants. All participants were in different locations, except for one desktop participant who was in the same physical location with one HMD participant. The desktop participant also acted as an on-site assistant to ensure physical safety of the HMD user in the same location (Figure 5).

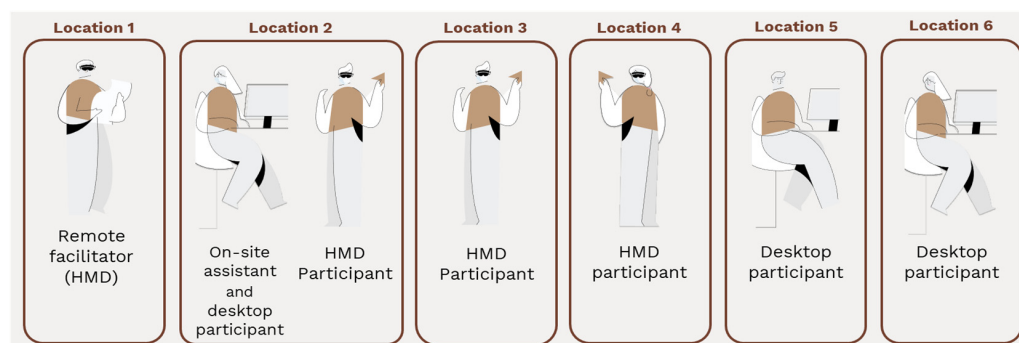


Figure 5. User test setup in Case 2.

Case-specific challenges: Because of restrictions on entering the offices, it was difficult to recruit industrial subject matter experts. In order to involve necessary subject matter expertise, one of the HMD participants was sent a VR set to their home and was instructed remotely via Teams on its use. The participant had no previous experience of setting up

a VR environment (e.g., calibrating the play area and updating the VR software). The participant was ~35 years old, with expertise in industrial maintenance, maintenance method development, and hands-on expertise from the field.

3.3. Case 3: VR for Maintenance Method Development and Documentation Creation

Case description: The aim of the study was to analyze the subjective perceptions and usefulness of a single-user VR system in accordance with maintenance and technical documentation industrial tasks. The study was designed in collaboration between academic and industry researchers. The test sessions were conducted at the company's premises in Finland and were facilitated and observed by industry researchers. In total, there were one pre-test, six on-site user tests, and one remote user test facilitated for a participant located in the USA.

Pandemic restrictions: At the time of the evaluation, the Finnish government recommended adhering to remote work practices [56]. Hence, the evaluation was designed to ease remote observation. The user test setting consisted of cameras located in the VR room, a speaker, and a video live-stream via Teams.

Test setup: The evaluation was conducted using a mix of remote observation and on-site facilitation. The remote observation was conducted using a Teams meeting. For the pre-test and six on-site user tests, there were one facilitator and one participant in the VR room (Figure 6). However, the evaluation for the participant in the USA was conducted entirely remotely. For each session, at least one remote observer was present in the Teams call. The audio in the session was transmitted via Teams using a good quality speakerphone in the VR room. The on-site facilitator set up a USB camera to live-stream the whole procedure, from set-in up the VR room until end of the user test session. In addition, during the VR evaluation, the participants' VR point-of-view was live-streamed simultaneously as the VR room camera view. The Teams call was recorded to capture the facilitation and participant's actions for later analysis. Lastly, a GoPro Hero 3 camera was set up to record the room setup as an offline backup (Figure 7). The test procedure was conducted using an HTC Vive Pro headset with hand controllers.

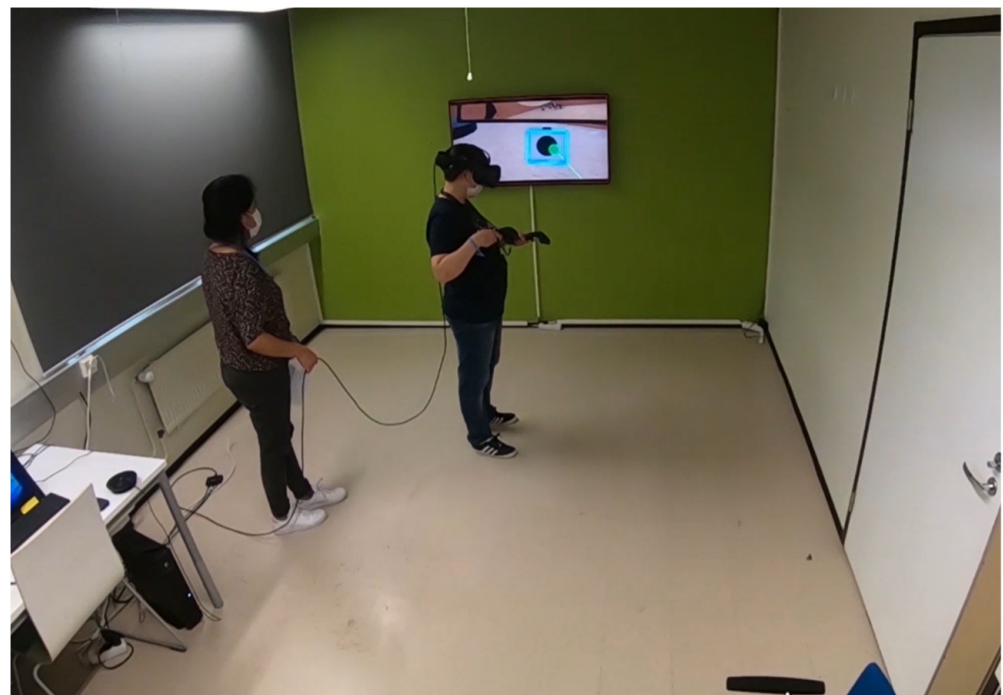


Figure 6. VR session, view from off-line camera during user tests in Case 3.

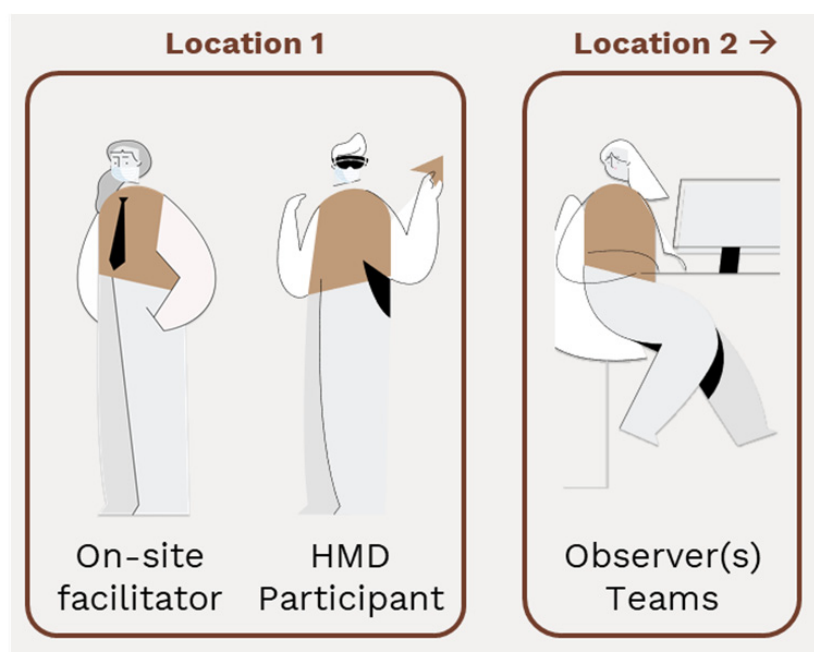


Figure 7. Test setup used in Case 3.

Case-specific challenges: The hardware setup was more troublesome than expected. During the first evaluation, the VR point-of-view was streamed over Teams, however, the VR room view was not. This hindered the remote observation as observers could not see the user's actions outside of VR. For that reason, the cameras and audio setup were changed to afford the VR room view simultaneously with the VR point-of-view (Figure 8). In addition, since observers were set up remotely, the on-site facilitator's duties became more demanding. The facilitators disinfected, set up, facilitated, and managed after-test practicalities by themselves. In practice, this meant that even though the evaluations lasted on average close to two and half hours, the facilitator needed one hour for preparation and two for after-test practicalities. There were network issues during the user tests which affected the observation. In Figure 9, the facilitator is visible on the screen twice, and the image is pixelated.

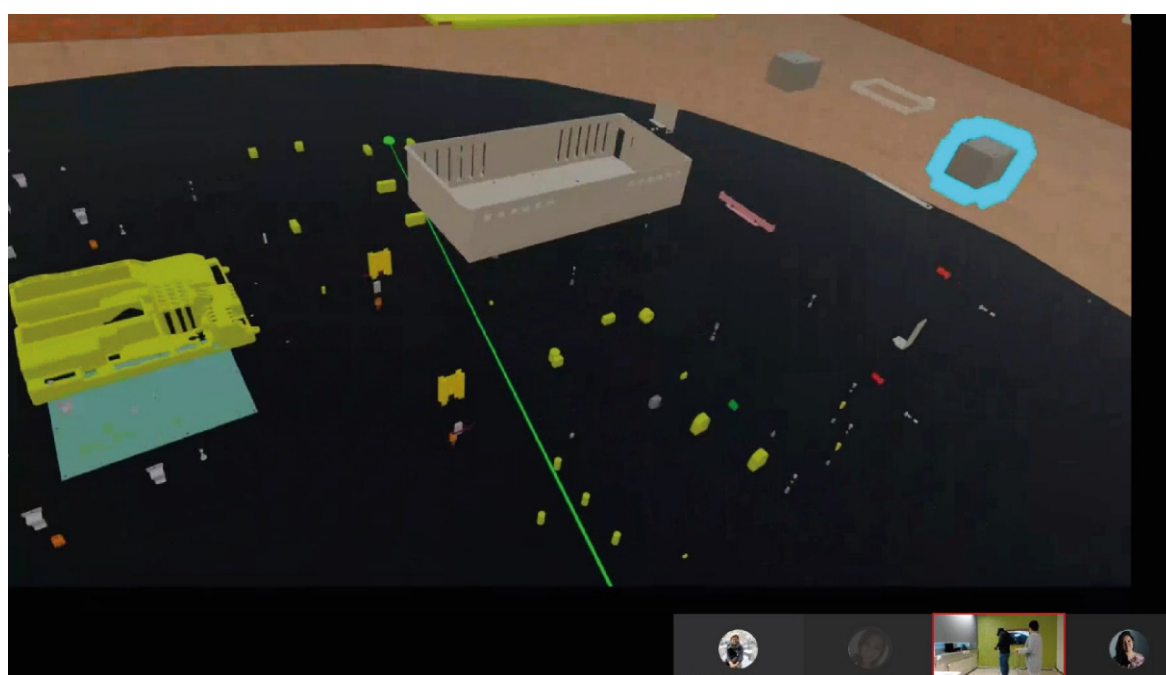


Figure 8. Single user evaluation observed via Teams. The HMD participant's view and USB camera view of the testing room were streamed for remote observers in Case 3.

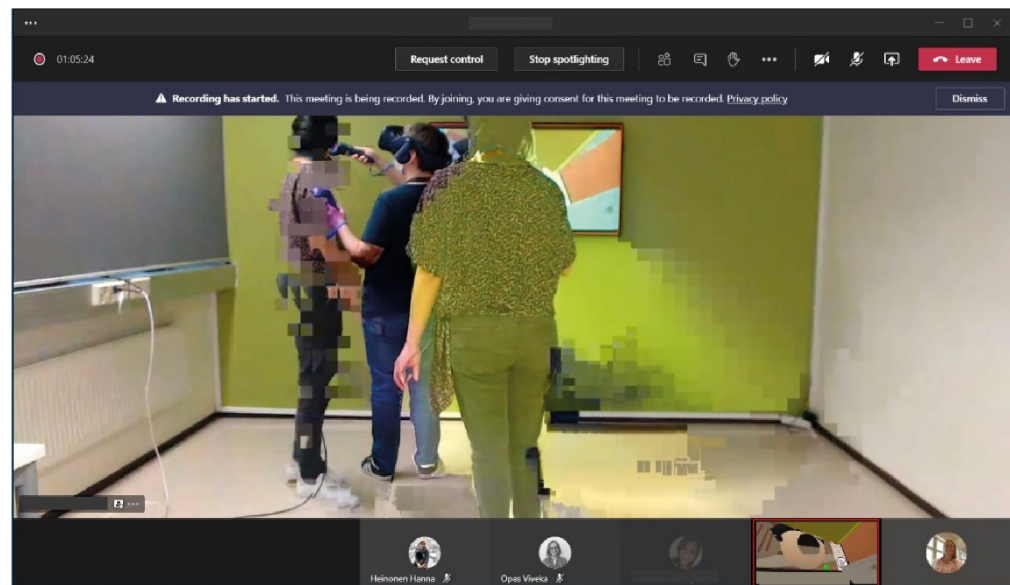


Figure 9. Network issues during remote observation in Case 3. The facilitator is seen twice in the image.

3.4. Case 4: Multi-User VR for Collaboration in Complex Machine Room Planning

Case description: The use of a multi-user VR in complex elevator machine room planning of high-rise building projects was tested. The study aimed to find out the benefits of multi-user VR from the user and business perspective as well as how to optimally apply the technology in the process. Four real-life high-rise building projects in Indonesia, Malaysia, the United States, and the United Arab Emirates were studied. There were nine HMD participants who were members from the selected projects and company personnel from Finland. There were no participants from the Malaysian site because a COVID-19 national lockdown prevented them from accessing the VR facility. Three remote cross-country user tests were organized. Each test involved participants from two to three different countries. Facilitation was conducted remotely, and the remote facilitator was in the virtual environment with the HMD participants. After each test, a group interview was held to gather feedback from the participants. In this study, four VR headsets were used: the HTC VIVE, HTC VIVE Pro, Oculus Rift S, and HP Reverb.

Pandemic restrictions: Movement restrictions in all testing countries were in effect, which restricted the use of the office and the number of people in one meeting room. International travel was prohibited. Access to VR facilities was limited. Participants were required to wear a mask at all times and wash their hands with soap and use hand sanitizer regularly.

Test setup: Before every user test, remote onboarding sessions were held to instruct the on-site assistant and familiarize HMD participants with the VR environment. In each testing session, 2–3 testing sites in different countries were involved. Each site had 1–3 HMD participants and 1 on-site assistant (Figure 10). Only one HMD participant at each site was in the VR environment at a time. They took turns to perform the assigned tasks. The whole session was remotely facilitated, and the communication was done via Teams. Each testing session focused on one machine room project so that its participants could experience the use of VR in their actual workflow. The VR view of the HMD participant, who was working daily on the examined project, was shared on Teams. Each site also had their camera on in the Teams video call for remote observation (Figure 11). Moreover, the remote facilitator juggled being in the VR environment for any requested VR task-related assistance with observing the HMD participants through their video on Teams. Furthermore, each testing session also had 1–2 desktop participants and remote observers who joined the online Teams meeting (Figure 12). The desktop participants did not perform many tasks for the testing but rather assisted with elevator technology related questions.

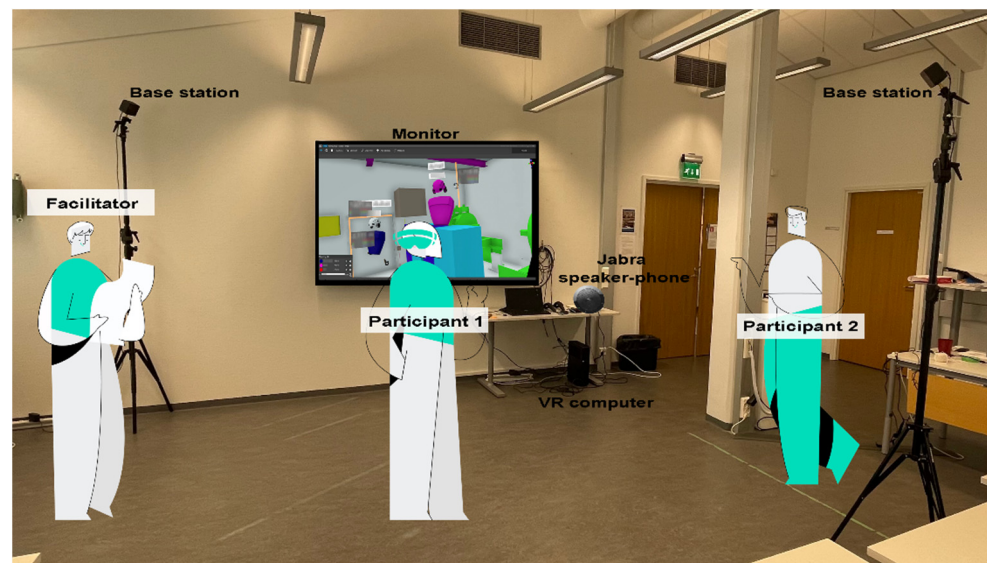


Figure 10. The physical setup of each testing site in Case 4. The base stations were only needed with HTC Vive, not with With Oculus Rift S.

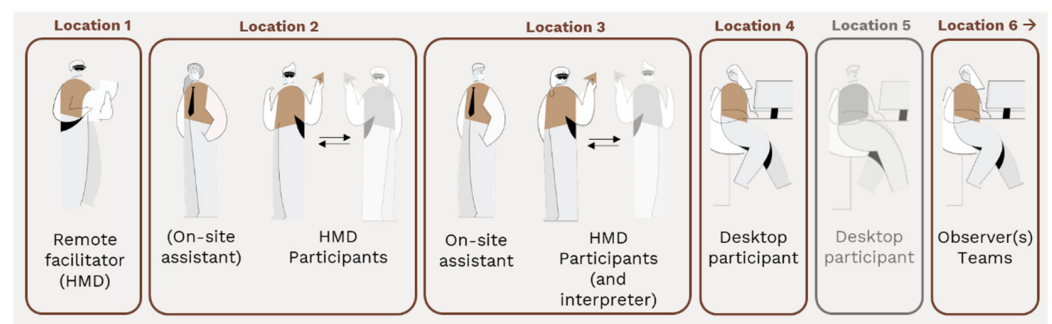


Figure 11. The user test setup in Case 4.



Figure 12. Screenshot of Teams view of one user test session in Case 4, showing HMD participants' view and USB camera video views from two testing locations.

In one instance of user testing, the on-site assistant was not available in one of the testing locations. The number of HMD participants was 1–3 for each testing location. The HMD participants took turns, and others participated via a large display on the wall. In one of the sessions, one HMD user also acted as an interpreter.

Case-specific challenges: Remote coordination between two different testing sites in one session was challenging since the remote facilitator did not have control over the test arrangement. The remote facilitator juggled being inside VR and observing what was occurring on Teams, making the facilitation demanding. Moreover, the on-site assistant was not available for one of the sessions. This added to the workload of the remote facilitator, who had to perform troubleshooting when technical difficulties occurred. Moreover, an unstable internet connection affected the synchronization to the cloud-based VR environment, causing differences in what users in different countries saw (Figure 13). This issue also created a delay in the communication on Teams. The situation became more challenging due to the language barrier. One user test was also cancelled due to a total lockdown in that country.



Figure 13. A VR scene of components falling apart due to a synchronization issue, Case 4.

3.5. Case 5: Collaborative Multi-User VR for Maintenance Method Development and Documentation Creation

Case description: In this study, we analyzed a collaborative multi-user virtual environment for maintenance method development and documentation creation. The aim of the study was to analyze the subjective perceptions and usefulness of a multi-user asymmetric VR system in accordance with industrial tasks. Asymmetric VR settings refer to the use of immersive and non-immersive platforms simultaneously, in this case by HMD and Teams participants. The study consisted of five user test sessions. For each session, two HMD participants and two Teams participants were present, with 20 participants in total. The two HMD participants joined the virtual environment simultaneously while the two Teams participants, from India, China, or the USA, joined the session using a Teams call (Figure 14). Teams participants engaged in the VR session by viewing the VR session through one of the HMD participants' point-of-view, referred to as HMDe, and interacted with voice commands via Teams (Figure 15). The test procedure was conducted using two HTC Vive Pro headsets with hand controllers.

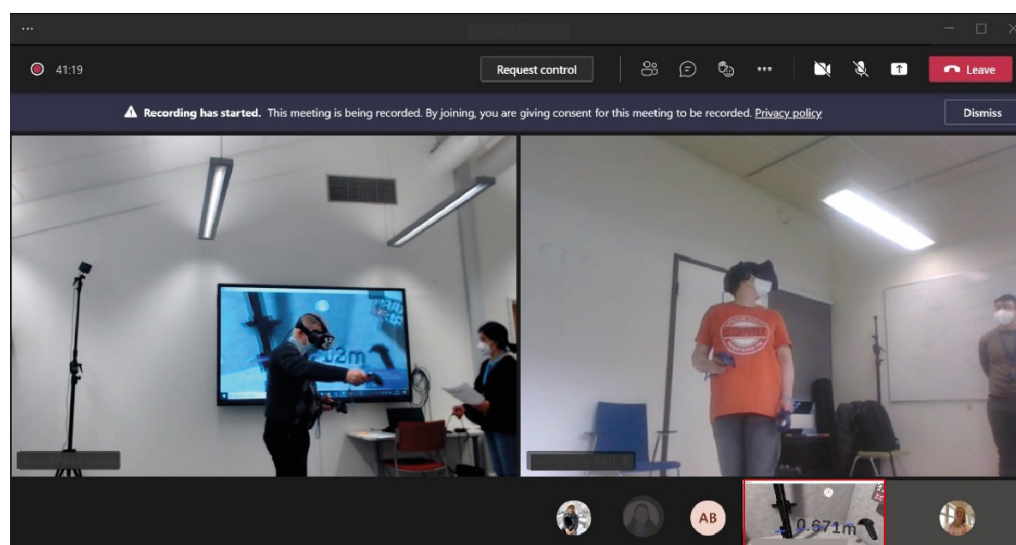


Figure 14. Multi-user VR session using Teams in Case 5. Two VR rooms are shown as the large images, and the HMDe view is in the small image on the bottom right.



Figure 15. An HMDe participant's point-of view and the two VR rooms shown via Teams (small images on the bottom left).

Pandemic restrictions: At the time of the evaluation, the Finnish government recommended adhering to remote work practices [56]. In addition, there were similar restrictions in place in India and the USA. The guidelines in China allowed the participants to work at the company facilities.

Test setup: Due to the COVID-19 situation, the sessions were designed to facilitate remote participation. For each session, two on-site HMD and two Teams participants were present. The facilitation consisted of two on-site facilitators, one chat facilitator, and remote observers. For the on-site sessions, there was one on-site facilitator and one HMD participant per location (See Figure 16).

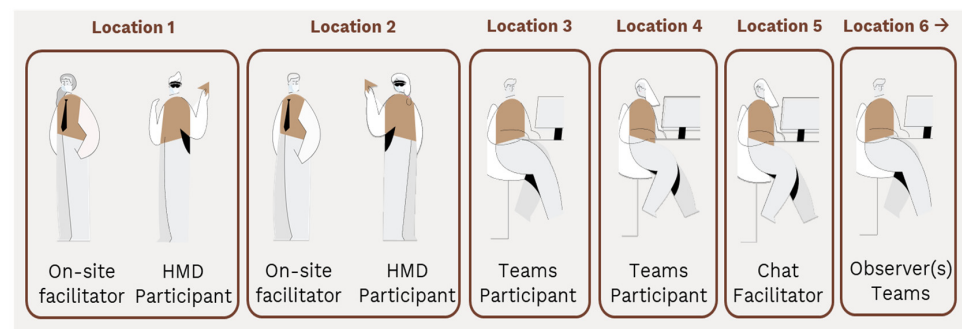


Figure 16. Test setup of Case 5.

The on-site facilitator kept two meters distance from the HMD participant during the whole evaluation. The invitations for the on-site sessions (HMD participants) included the COVID-19 safety guidelines and required participants to read and agree with the participation guidelines upon acceptance of the invite. To facilitate remote participation, a series of cameras were placed in the two on-site testing locations to transmit and record the session in Teams (Figures 14 and 15). In addition, the HMD's view was transmitted during the evaluation. The Teams session was recorded and stored for later analysis. The test procedure was conducted using two HTC Vive Pro headsets with hand controllers.

Case-specific challenges: This user test was conducted under one of the highest pandemic-restriction periods in Finland [56]. Thus, approval to conduct the test was required from the company's human resources and upper management from all sites involved in the test prior to the evaluation. Subsequently, the researchers corroborated the participants' right to decline their participation. One on-site participant declined and one Teams participant postponed their participation. Other technical practicalities of the evaluation regarded the general setup. For instance, the audio in one of the on-site evaluation rooms was set up using a speaker to accommodate the HMD participant and the on-site facilitator. Conversely, the audio from the other evaluation room was set up directly in the VR headset to afford only the HMD participant. Lastly, due to the varying roles of the remote participants, the facilitation had to be adjusted to suit the purpose. During the pre-test, we realized that the facilitation roles were cumbersome. Initially, only the two on-site facilitators were included in the evaluation script; however, with this setup, the Teams participants were left without technical support. For this reason, after the pre-test we decided to include a chat facilitator to supplement the remote facilitation and to provide technical support. In addition, we faced challenges communicating the varying range of roles and facilitators involved in the evaluation. This was mainly because the communication channels differed from on-site to remote facilitation; however, ultimately this evaluation included both types of facilitation at once.

3.6. Safety and Hygiene for On-Site VR Tests during the Pandemic

The safety protocols were enforced before, during, and after the user tests. Before the session, the facilitator and participants needed to wash their hands and use hand sanitizer before entering the test room. A sanitization table was set up at the entrance of the testing room, featuring FFP2 face masks, hand sanitizer, and surgical gloves for the participants and facilitators. People who could not wear a mask or use hand sanitizer did not participate in the test. In addition, a UVC light was used to disinfect the headset and hand controllers, and the equipment was wiped with disinfecting wipes. To ensure the sanitization of the devices, VR disposable hygiene covers were placed after disinfection of the equipment.

The facilitator and participant wore FFP2 face masks and surgical gloves throughout the duration of the user test. The participants were instructed on the use of the VR equipment with a training video along with instructions from the facilitator, so each participant put the VR equipment on by themselves. Due to the practicalities of VR facilitation, the protocols also included guidelines to prevent physical contact between

the facilitator and the participant. In case the participant was too close to a physical obstacle, the facilitator would warn the participant by saying “stop” and instruct in which direction to move. The facilitator was only advised to approach the participant in the case of emergency: e.g., if the participant feels dizzy and is at risk of falling/fainting. Fortunately, no such event occurred during the user tests. Lastly, remote participation was enabled, the testing session was streamed via a USB camera in the testing room using Teams. Participation consent forms were shared with the participants via Teams, and participants signed the form by adding their name to the Teams chat.

4. Methods

A mixed-method approach was used in the user tests, but it is not described here as it is not relevant for the results of this article. During each of the tests, research notes and pictures were taken, and videos were recorded. All the research materials were stored in company and university SharePoint databases. Safety and hygiene protocols used in user tests were recorded in PowerPoint files and stored in the company SharePoint database (see Supplementary Material S1).

After all the user tests were completed, we analyzed all the available research materials to collect learnings and insights related to organizing remote user tests. Each case review and analysis involved creating an affinity diagram from the notes. In addition, we had group discussions and workshops with researchers who were involved in the cases.

5. Results

The combined results of the five user tests can be categorized into the following groups: general user test arrangements, insights that are important because of new and fascinating technologies, insights that are important due to a pandemic, insights that are important because of the remote setup, insights that are important because of the industrial context, and insights regarding remote interviews. In this section we also give recommendations for safety and hygiene protocols for VR user testing.

5.1. Insights that are Especially Important Due to a Pandemic

In a pandemic, it is important to maintain social distancing (2 m), to disinfect devices and the premises between users, and to have enough time between sessions for ventilation (Case 3, Case 4, Case 5). Considering all preparations, disinfecting, and restrictions, in the ideal case only one test per day should be scheduled (Case 3). The room used for testing should be large enough to accommodate social distancing (Case 3, Case 4, Case 5).

Under normal conditions, the practice is that the on-site assistant or facilitator helps with putting on the VR equipment. Because of the 2 m distance rule, this is not possible during a pandemic, and therefore good instructions for (first-time) users are extremely important. Video instructions displayed on a large display proved to be efficient (Case 3, Case 4, Case 5).

Due to social distancing, travel limitations, and restricted gatherings and contact, it is important to enable remote participation and remote observation (Case 1, Case 2, Case 3, Case 4, Case 5). With VR testing, this is best supported with hybrid setups, where both the VR view and room view are shared over Teams (or another video conferencing app) for remote participants and observers (Case 1, Case 2, Case 3, Case 4, Case 5).

In Case 1 and Case 2, the participants were sent the VR set and a compatible laptop. What is remarkable is that these two participants were far from the tech-savvy type, both having a mechanical engineering and field maintenance background, with very little experience of VR. They had never set up VR glasses themselves, and only participated in some facilitated VR sessions before. Furthermore, the VR set they were using, the HP Reverb, was a new device for them. In these cases, they had a Teams video call with a VR expert. The VR expert went through the setup procedure step-by-step with them, and with the video connection the VR expert was able to see everything that happened at the remote location, for example, which lights were active on the hand controllers (Case 1, Case 2).

Due to the changing conditions and changing rules, the development was done in iterative sprints. Therefore, we had a slightly different setup in Case 1, which made comparisons difficult. This was the first case and the pandemic restrictions were new. It is advisable to plan carefully and be prepared for the worst-case scenario. In later cases, we planned the experiments so that changing restrictions did not have an effect on executing the user tests (Case 2, Case 3, Case 4, Case 5).

When facial masks are required, facial expressions are not fully visible to observers and, therefore, are difficult to interpret. Facial masks also muffle the voices of the participants. Therefore, good quality audio is essential. Furthermore, the think-aloud method is very important to understand the participants and their actions. When a mask was used, the verbal expression of feelings was became important in addition to thoughts, which was explained to participant. (Case 3, Case 4, Case 5).

5.2. Insights that are Especially Important Because of Remote and Hybrid Setups

It is good to have Teams or a similar video collaboration tool to support user test arrangements if done remotely, even if the tests are carried out in VR. It can be at least used in the orientation before logging in to VR (Case 1, Case 2, Case 3, Case 4, Case 5).

It is always advisable to record the user test session for later reference, but when observing user tests remotely, a live video feed is also essential. The integrated laptop camera is designed for “face-to-face” video conferencing, and it is difficult to position the laptop so that the user test environment and participants can be seen from that camera feed. Therefore, we have noted that it is useful to have an external USB camera, which can be easily positioned to cover the whole user testing area (Case 1, Case 3, Case 4, Case 5). An additional USB webcam is also useful if there is a need to introduce the use of new equipment, or if there are some practical issues like setting up the VR equipment for participants (Case 1, Case 2).

In Case 1, we asked the participants to fill in a questionnaire after the session. One person never filled in the questionnaire, and as it was anonymous, we did not know whom to remind. Thereafter, in all other user tests, we closed the session only after the participants had answered the questionnaire, and we could be sure that everyone filled out the questionnaire without any problems (Case 3, Case 4, Case 5). Thus, we recommend this practice.

In some user test sessions, the participant was alone in a remote location (Case 1, Case 2, Case 4, Case 5). This sometimes led to a chaotic situation without the on-site assistant (Case 4). Hence, it is recommended to have at least one on-site assistant at each testing location, if possible (Case 1, Case 4).

Moreover, the physical view and VR view of each testing site should be enabled. This helps the remote facilitator comprehend the circumstances of the testing site and determine the exact guidance needed for all questions raised. This is especially important if there is no on-site facilitator in the remote location. The remote observers benefit from multiple views as well (Case 3, Case 4, Case 5).

Instructions (a comprehensive user test plan) should be always available and communicated well in advance. Everybody can then monitor the progress and prepare for their upcoming tasks. This is especially important in the remote setup because of limited communication possibilities during sessions. If the researchers are physically in the same room, one may whisper and discretely ask something from a colleague without disturbing the session (Case 1, Case 3, Case 4, Case 5).

It is preferable to have instructions easily available for on-site participants and directly in the Teams chat box for remote participants. This approach provides immediate access and eliminates the need for additional hardware and software manipulation (Case 1, Case 3, Case 4, Case 5).

If the facilitation is done remotely, the setup needs to be very carefully tested. A full pre-test session or “a dress rehearsal” is recommended where the whole test workflow with all the roles is tested. The test setup needs to be identical with the actual user test to discover any issues with the setup or workflow. We also recommend having someone with a technical background on site for at least the pre-test to quickly solve any technical issues with, for example, connectivity (Case 2, Case 3, Case 5, Case 5).

The use of face masks impedes the observation of participants’ facial expressions and muffles their voices. This is especially poignant when using the think-aloud method and observing remotely, and the camera angle is not always optimal. Therefore, we recommend using several cameras, if possible, to overcome this and instructing the users to speak clearly (Case 3, Case 4, Case 5). The audio and camera settings and the network should be tested during the pre-test and prior to every user test with the remote observers (Case 3, Case 4).

The peculiarities of the remote observers’ computer setup, internet connection, software knowledge, and other variables, might affect the evaluation setup. If the roles and responses regarding technical support during the test session are not clear, it could cause confusion. Furthermore, when someone starts to solve technical issues, they are not able to concentrate on their actual role, e.g., facilitation/observation (Case 3, Case 4). On the other hand, when there is a dedicated person to help with the technical issues and the software being used, everyone is able to concentrate on the actual task and their own roles (Case 1, Case 2, Case 4). Therefore, it is advisable to have an extra technical support facilitator participating in the user test sessions if possible, or at least have roles and responsibilities agreed upon for such situations where technical support might be needed (Case 1, Case 3, Case 4).

In the case of hybrid on-site and remote facilitation, the evaluation procedure should be supported with online tools. For instance, the VR tasks were administered as notes in the VR environment so that HMD and non-HMD participants could read them. However, the remote tasks were administered by the chat facilitator in the chat window following the task list. In addition, the whole procedure was supplemented with an online presentation shared over Teams. The slide deck followed the sections described in the user test plan and helped all participants have a visual example of the upcoming items in the evaluation (Case 5).

In addition, at the end of 2020, Teams added an automatic transcription for recordings which aided in note taking (Case 4 and Case 5).

5.3. Insights that are Especially Important Because of the Industrial Context

In the industrial context, some on-site facilitators and assistants are company employees, not researchers. Therefore, they do not necessarily have experience with facilitation or research methodology. Thus, there is a strong need for detailed instructions and a pre-test session. In particular, the role of the facilitator and on-site assistant needs to be clearly defined, communicated, and tested in the pre-test session. Otherwise, this might lead to the on-site assistant taking a stronger role or misunderstanding their role and the remote facilitator losing control over the test environment and facilitation (Case 4). Similarly, if the interpreter is not a professional interpreter, the role needs to be made clear so that the opinions and the views of the interpreter are not mixed with those of the participants (Case 4).

Recruiting participants is trickier in the industrial context, where certain expertise is needed, than, for example, in contexts where general usability of an application is being tested. In the latter case, the background of the participant does not matter that much, and it is common practice in academia to recruit students from campus. In our industrial case, it was important to have real subject matter experts attending the user tests. However, due to the pandemic restrictions, some were not able to participate (Case 3 and Case 4). In Case 1 and Case 2, we enabled the participation of important subject matter experts by sending VR devices to their home and instructing them remotely on the setup.

In industrial research, it is important that tasks in the user test mimic the actual work or, even better, real tasks. This way, the participants understand the need for research input as it is directly related to their own work. Furthermore, this way the collected feedback is focused on improving the industrial work.

A good introduction to user testing is needed as people often come in a hurry from other work duties. Therefore, we have learned that a good practice is to have refreshments available at the beginning and allow participants to relax and have a small break while explaining the purpose and context to get them onboarded into testing. Again, this is very different from recruiting students from campus. Of course, remote setups made it difficult to offer refreshments, but in some locations the remote assistant was able to organize it.

5.4. Insights Regarding Conducting Remote and Hybrid Interviews

In a hybrid interview setting, many factors could affect how interviewees behave in the session. The different nature of being remote and on-site is among the most influential. On-site interviewees tend to dominate the discussion as the interview is conducted in a face-to-face format. Remote interviewees might find it hard to keep up with the conversation. Hence, they may find that their role is less significant in the discussion. Moreover, cultural and personal differences may also affect how active people are in the remote interview. Some are more open to discussion on-site while some might find it easier to communicate remotely. The language barrier could also have a significant impact as some may not be able to keep up with the discussion in a foreign language (Case 4, Case 5). Therefore, it is important that the interviewer proactively engages everyone throughout the session. Using a non-professional interpreter may lead to a situation where the interpreter dominates the discussion over the participants (Case 4).

Every interviewee should be given a chance to speak and asked or encouraged by the interviewer to express their opinion in all questions. One may hesitate to answer as they deem their responses repetitive to others, which might not be the case. Tackling this assumption by asking everyone to express their views can result in a more fruitful discussion. In addition, the interviewer can address a question to a specific participant to initiate the conversation (Case 4, Case 5).

The note-taking procedure is usually cumbersome during in-person interviews as interviewers usually rely on physical notes. In contrast, remote interviews improved the review and note-taking procedure during the interviews (Case 4, Case 5). This is because in remote interviews, interviewers had easy access to digital notes. In addition, remote interviews over platforms such as Teams allowed the interviewers to have access to multiple monitors. Essentially, the interviewer set up the online interview on one screen and the notes on another.

In addition, video conference tools expanded the ways to share ideas. In some cases, the interviewers and interviewees used the chat box to type the discussion (Case 4, Case 5). This was especially observed in interviews where the participants preferred and felt more confident writing rather than speaking in a foreign language. Similarly, video conference software enabled the interviewers and interviewees to reinforce the communication using visual tools. For instance, some interviewees used the screen-sharing tool to share slide decks or images to support the discussion (Case 3, Case 4, Case 5). In other cases, the interview was supplemented with applications, such as Miro, to brainstorm and describe new ideas (Case 4).

In all interview sessions it was observed that the use of cameras to see the facial expressions enhanced the engagement in the conversation. Likewise, seeing the facial and hand expressions eased the interpretation of the discussion in later analysis. The use of a camera was particularly beneficial in the case of group interviews where the interviewees reflected on their colleagues' answers. In addition, all the interviews benefited from seeing the interviewer and other participants on the live stream (Case 4, Case 5).

5.5. Recommendations for On-Site Safety and Hygiene Protocols for VR User Testing

On-site safety and hygiene protocols provided comfort to on-site facilitators and participants as they were assured that the researchers were compliant with the local regulations. Similarly, the researchers adhered to company regulations and thus worked closely with managers and human resource representatives. Clear and concise safety and

hygiene guidelines helped us to communicate safety protocols and demonstrate to the company authorities and participants that the user tests would follow official regulations and guidelines.

As the pandemic situation and official guidelines changed, we had to iteratively change our procedures. Figure 17 presents the ultimate on-site safety measures that we used during the pandemic in VR user tests. Furthermore, the safety measures taken by the on-site facilitators and participants are presented in Figure 18.

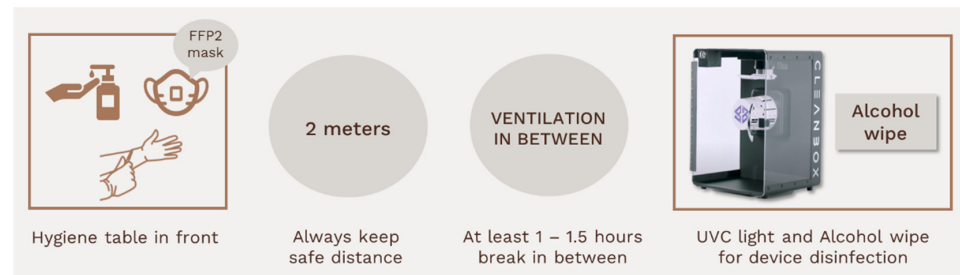


Figure 17. On-site safety measures for organizing user test during the pandemic.



Figure 18. Safety measures taken by the on-site facilitators and participants throughout the user testing during the pandemic.

5.6. Insights That Are Especially Important Because of Novel and Fascinating Technologies

When you are testing novel, exciting technologies, such as VR, test participants might get excited and attempt things outside of the scope of the test plan. In Case 2, we ran out of time due to this and did not collect all the feedback during the session as planned. Therefore, some of the feedback was collected afterwards via email. Our recommendation is that if there are people who have not used the equipment or technology before, reserve enough time for users to explore the environment and its features before the actual user test. The same recommendation applies to all exciting technologies (Case 4, Case 5).

As mentioned earlier, it is sometimes advisable to have a separate introduction and familiarization session before the user test.

5.7. General Comments Regarding User Test Arrangements

Especially in the case of complex evaluation setups, such as Case 5 with the combined in-person and remote evaluation, it is imperative to script a comprehensive user test plan. In the plan, all the evaluation steps should be clearly specified, and action points appointed to specific people: participants, facilitators, and observers. If possible, the script should also include estimate times per section so that the evaluation remains on track. Prior to the evaluation, the facilitators, assistants, and observers should have a debrief meeting to overview the test plan. During the evaluation, the facilitators should ensure that the participants comprehend the procedure, roles, and tasks.

User test sessions should be scheduled in a staggered manner, with preferably only one or two tests per day. This will give the facilitator and observers time to finalize notes and prepare for the next tests in an orderly manner. This is especially important when the facilitators and observers have other work duties alongside the user tests, as is often the case in industry research activities. This is also important during a pandemic to have enough ventilation time for test facilities between test sessions, and to reserve time for disinfecting procedures.

When conducting Case 1, we had to adapt to the changing situations and setups on an ad hoc basis, which made the tests challenging. Therefore, a pre-test practice was used for Cases 2, Case 3, Case 4, and Case 5. The pre-test worked as a “dress rehearsal” where any complications and difficulties were noticed, and no ad hoc fixes or adjustments were needed in the actual user tests. In most cases, one pre-test was enough to finetune testing procedures, but in Case 5, for example, several pre-tests were needed before the setup was clear and manageable.

Recording the user test sessions is highly recommended. Researchers can later conduct further analysis of the verbal and non-verbal feedback from the users, which might not be thorough while facilitating the session. The audio and camera settings should be tested during the pre-test and prior to every session.

It is a good practice to have a video conferencing connection with recording (e.g., Teams session) of all organizers and participants throughout the whole session, from the beginning until submission of questionnaires or whatever is the last part of the test session. This ensures that everyone is part of the whole end-to-end process and as much information as possible is gathered (Case 1, Case 2, Case 3, Case 4, Case 5). For research and note taking purposes, it is good to also record the preparations (e.g., setting up VR gear and cameras) (Case 3, Case 4, Case 5).

In earlier work, we had introduction + user test + interview sessions that took close to 2.5 h, and it proved to be too long and intensive for the users. We noticed that a separate introduction and training session reduced the intensity of actual user testing and allowed more time to be spent on each task and on discussion arising from the think-aloud method during the session. In Case 1 and Case 4, we had a separate 0.5–1.5 h introduction and training session for the VR. Thereafter, within 1–3 days, we had the actual VR collaboration user test with work-role-related tasks. In both cases, the user test and discussion lasted 1.5–2 h. However, in Case 4, each user spent only 20–30 min in an intensive VR environment as their role was rotated. In Case 5, the training session and user test were held in succession, and the whole session with the training, user test, surveys (before and after the user test), and group interview took up to 2 h and 45 min, but it included a break and the schedule was not tight. As our cases were to test the industrial use of collaborative VR rather than general usability, familiarization helped users to shift their focus and provide more relevant comments on their work role and tasks. Such results were achieved with relaxed scheduling throughout all cases.

6. Discussion

Even though arranging user tests during a pandemic with all the restrictions is challenging, with good planning and proper setups it is possible, as this article proves. Yet there are some differences and limitations in remote user testing compared to face-to-face testing sessions. For example, on-site people often continue discussion in an informal manner after finishing the user test tasks. This happened, for example, in Case 3 and Case 5 with some on-site participants. In many remote setups, possibility for this kind of discussion is missing and valuable comments are lost.

Some people are shy and/or anxious to speak when several people are listening, and prefer to share their views in one-to-one discussions. Some participants also noted that they feel nervous because of the remote observers (Case 3, Case 5). Thus, remote observers should remain as unobtrusive as possible, so the participant acts as natural as possible (see, e.g., [57]). We also highlight the importance of good facilitation and encourage getting everyone deeply involved.

In remote and mixed setups, it is important to have clear roles for organizers and to have enough researchers involved. In the ideal case, there is a dedicated person for resolving technical problems. An on-site facilitator should be able to concentrate on the persons that are physically located at the site; therefore, it is advisable to have a dedicated chat facilitator, or similar role, to narrate and facilitate remote participants.

In an on-site session, it is natural and easy for observers to move a little when needed to have a better view of the participants during the session. In remote observation this is not possible. Thus, we recommend having several cameras, a movable camera, and an onsite assistant who can move the camera. In addition, there is always a risk of network issues during remote sessions, and, therefore, a recording of the session and an offline camera as backup are encouraged.

It is always important to plan user tests properly, but when implementing a remote user test, the planning becomes even more important. Ad hoc changes are easier to manage in on-site sessions. The remote setup is especially demanding when several remote locations are involved, and when non-research oriented remote assistants are used. Thus, good planning with back-up plans for possible incidents is needed.

One good practice is to have a pre-test (“dress rehearsal”) practice where the whole user test procedure and setup is tested. The roles of all participants should be clear and well defined. This is especially important when utilizing industrial employees as on-site assistants or interpreters.

During a pandemic, the unexpected sometimes happens: new restrictions are announced, people are quarantined, or for some reason on-site participation is cancelled at the last minute (Case 4, Case 5). Location-independent remote participation reduces the risk of cancellations and rescheduling. In Case 1 and Case 2, the test devices (VR equipment) were sent to some participants’ homes to enable their participation.

The role of hygiene and safety in user testing during a pandemic cannot be stressed enough. During more normal times, the facilitator is responsible for safe working practices, such as making sure that the participants wear the required personal protective equipment (PPE). During a pandemic, the role of safety is increased with requirements for wearing masks, disinfecting the equipment, social distancing, etc. This might be challenging remotely and requires careful planning. Additionally, if the remote participant is alone in the remote location, physical safety must be taken care of for the HMD participant. Researchers [30,31] have noted that a video-based approach enables the researcher to understand exactly the steps of the participant during the evaluation. Thus, having a camera view of the remote location is highly recommended in remote and mixed (on-site and remote) evaluations (Case 3, Case 4, Case 5).

Even though the pandemic has made many areas of life challenging, there are also good practices that we have learned that can be used to enhance post-pandemic global collaboration. We started planning Case 1 in the pre-pandemic era, but as the pandemic hit, we adjusted our plans to accommodate for the new conditions. With learnings from each of the cases, our processes and the execution of the remote and hybrid user tests became smoother, and many of these ideas can be applied to the post-pandemic world. By working in remote and hybrid setups, we can involve a global user base in the testing, save on travel money, and encourage sustainability.

7. Conclusions

First of all, it is important to plan very carefully when organizing hybrid user tests. To support this, we created a checklist for organizing VR user tests in hybrid (on-site and remote) setups (Supplementary Material S2). As these kinds of setups, especially VR tests, require many researchers and assistants, it is important to have clear roles and responsibilities. We recommend a full pre-test session before starting the actual tests. The use of several cameras and back-up plans for network issues are recommended. To support the practical arrangements of hybrid VR user tests, we collected a list of tips for organizers (Supplementary Material S3).

During the pandemic, increased hygiene and safety protocols were implemented. It is recommended to continue these established practices also in the post-pandemic user tests to prevent, for example, milder flus (Supplementary Material S1).

Communication in a shared physical space is natural. When communication is done remotely, it is important to pay more attention to establishing human-to-human connection: keep cameras on, encourage think-aloud, and facilitate so that everyone is actively involved.

On-site user tests cannot always be totally replaced with remote user testing, but it should be considered as an alternative. On many occasions, it is possible to get answers to research questions also with a remote test setup: if not to the full extent, still sufficiently to continue the research. On the other hand, remote user testing can often fully support the work. Reasons to prefer remote testing are to have faster testing cycles, to involve more participants from different cultures or different locations, and sustainability, among others. When traveling is not possible or preferred for one reason or other, with careful planning the research can continue with remote setups: “Where there’s a will, there’s a way”.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/mti5100062/s1>, Supplementary Material S1: On-site safety and hygienic procedures for VR user testing, Supplementary Material S2: Checklist for organizing VR user tests in hybrid setups, Supplementary Material S3: Tips for organizing VR user tests in hybrid setups.

Author Contributions: Conceptualization, S.S., H.H., A.B., P.B.P. and P.T.; methodology, S.S., H.H., A.B., P.B.P. and P.T.; software, validation, formal analysis, S.S., H.H., A.B., P.B.P. and P.T.; investigation, S.S., H.H., A.B., P.B.P., P.T. and V.O.; resources, data curation, writing—original draft preparation, S.S., H.H., A.B., P.B.P., P.T. and V.O.; writing—review and editing, S.S., H.H., A.B., P.B.P., P.T. and M.T.; visualization, P.T. and P.B.P.; supervision, S.S. and M.T.; project administration, S.S. and M.T.; funding acquisition, S.S. and M.T. All authors have read and agreed to the published version of the manuscript.

Funding: This paper was partly funded by the Business Finland Human Optimized XR (HUMOR) co-innovation project [50], and partly funded by the Business Finland KEKO Smart building ecosystem project [51].

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Acknowledgments: We thank Alberto Gabriel for his work on Case 1, Matti Pouhakka and Jani Leskinen for their help with the Glue environment, and Iiro Naamanka for his help with the Design Space environment. Furthermore, we thank John Mäkelä, Jaakko Hakulinen, and Kimmo Ronkainen for their contribution to the HUMOR project and implementation of the COVE-VR environment. We also thank Peter Eagling, Daniel Maughan, and Amelia Veronica for their input and help in Case 4.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fernandes, N. Economic Effects of Coronavirus Outbreak (COVID-19) on the World Economy. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3557504 (accessed on 2 May 2021). [CrossRef]
2. Helm, D. The Environmental Impacts of the Coronavirus. *Environ. Resour. Econ.* **2020**, *76*, 21–38. [CrossRef]
3. Ratcliffe, J.; Soave, F.; Hoover, M.; Ortega, F.R.; Bryan-Kinns, N.; Tokarchuk, L.; Farkhatdinov, I. Remote XR Studies: Exploring Three Key Challenges of Remote XR Experimentation. In Proceedings of the Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems (CHI EA '21), Association for Computing Machinery, New York, NY, USA, 8 May 2021; Volume 121, pp. 1–4. [CrossRef]
4. Singh, R.P.; Javaid, M.; Kataria, R.; Tyagi, M.; Haleem, A.; Suman, R. Significant Applications of Virtual Reality for COVID-19 Pandemic. *Diabetes Metab. Syndr. Clin. Res. Rev.* **2020**, *14*, 661–664. Available online: <https://www.sciencedirect.com/science/article/pii/S1871402120301302> (accessed on 2 May 2021). [CrossRef] [PubMed]
5. User Research Community. Doing User Research during Coronavirus (COVID-19): Choosing Face to Face or Remote Research. Available online: <https://www.gov.uk/service-manual/user-research/doing-user-research-during-coronavirus-covid-19-choosing-face-to-face-or-remote-research> (accessed on 2 May 2021).
6. Suh, A.; Prophet, J. The State of Immersive Technology Research: A Literature Analysis. *Comput. Hum. Behav.* **2018**, *86*, 77–90. [CrossRef]
7. Ratcliffe, J.; Soave, F.; Bryan-Kinns, N.; Tokarchuk, L.; Farkhatdinov, I. Extended Reality (XR) Remote Research: A Survey of Drawbacks and Opportunities. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, Yokohama, Japan, 8–13 May 2021; Association for Computing Machinery: New York, NY, USA, 2021; Volume 527, pp. 1–13. [CrossRef]
8. Woodward, C.; Lahti, J.; Rönkkö, J.; Honkamäa, P.; Hakkarainen, M.; Jäppinen, J.; Rainio, K.; Siltanen, S.; Hyvärkä, J. Virtual and Augmented Reality in the Digitalo Building Project. *Int. J. Des. Sci. Technol.* **2007**, *14*, 23–40.
9. Gu, N.; Watanabe, S.; Erhan, H.; Haeusler, M.H. Immersive Virtual Environments: Experiments on Impacting Design and Human Building Interaction. In Proceedings of the 19th International Conference on Computer Aided Architectural Design Research in Asia, Kyoto, Japan, 14–16 May 2014; pp. 729–738.
10. Zaker, R.; Coloma, E. Virtual Reality-Integrated Workflow in BIM-enabled Projects Collaboration and Design Review: A Case Study. *Vis. Eng.* **2018**, *6*, 4. [CrossRef]
11. Havard, V.; Trigunayat, A.; Richard, K.; Baudry, D. Collaborative Virtual Reality Decision Tool for Planning Industrial Shop Floor Layouts. *Procedia CIRP* **2019**, *81*, 1295–1300. [CrossRef]
12. Schwarz, S.; Regal, G.; Kempf, M.; Schatz, R. Learning Success in Immersive Virtual Reality Training Environments: Practical Evidence from Automotive Assembly. In Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society (NordiCHI '20), Tallinn, Estonia, 25–29 October 2020; Association for Computing Machinery: New York, NY, USA, 2020; Volume 50, pp. 1–11. [CrossRef]
13. Guo, Z.; Zhou, D.; Zhou, Q.; Zhang, X.; Geng, J.; Zeng, S.; Lv, C.; Hao, A. Applications of Virtual Reality in Maintenance During the Industrial Product Lifecycle: A Systematic Review. *J. Manuf. Syst.* **2020**, *56*, 525–538. [CrossRef]
14. Schina, L.; Lazoi, M.; Lombardo, R.; Corallo, A. Virtual Reality for Product Development in Manufacturing Industries. In *Augmented Reality, Virtual Reality, and Computer Graphics, AVR 2016*; Lecture Notes in Computer Science; Springer: Cham, Germany, 2016; Volume 9768. [CrossRef]
15. Berg, L.P.; Vance, J.M. An Industry Case Study: Investigating Early Design Decision Making in Virtual Reality. *ASME J. Comput. Inf. Sci. Eng.* **2017**, *17*, 011001. [CrossRef]
16. Maher, M.L. Designers and Collaborative Virtual Environments. In *Collaborative Design in Virtual Environments*; Springer Netherlands: Dordrecht, The Netherlands, 2011; pp. 3–15. [CrossRef]
17. He, Z. Enhancing Collaboration and Productivity for Virtual and Augmented Reality. Ph.D. Thesis, New York University, New York, NY, USA, 2021.
18. KONE Corporation. Available online: <https://www.kone.com/en/> (accessed on 15 August 2021).
19. Radecki, J.; Schonfeld, R.C. The Impacts of COVID-19 on the Research Enterprise: A Landscape Review. Available online: <https://sr.ithaka.org/wp-content/uploads/2020/10/SR-Report-Impacts-of-COVID-19-on-the-Research-Enterprise102620-1.pdf> (accessed on 2 May 2021). [CrossRef]
20. Vitae 2021 Careers Research and Advisory Centre (CRAC). The Impact of the COVID-19 Pandemic on Researchers in Universities and Research Institutes. Available online: <https://www.vitae.ac.uk/impact-and-evaluation/covid-19-impact-on-researchers> (accessed on 2 May 2021).
21. Woolston, C. Pandemic Darkens Postdocs' Work and Career Hopes. *Nature* **2020**, *585*, 309–312. [CrossRef]
22. Rex Hartson, H.; Castillo, J.C.; Kelso, J.; Neale, W.C. Remote Evaluation: The Network as an Extension of the Usability Laboratory. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vancouver, BC, Canada, 22–27 April 2006; Association for Computing Machinery: New York, NY, USA, 1996; pp. 228–235. [CrossRef]
23. Castillo, J.C.; Hartson, H.R.; Hix, D. Remote Usability Evaluation: Can Users Report Their Own Critical Incidents? In *CHI 98 Conference Summary on Human Factors in Computing Systems (CHI '98)*; Association for Computing Machinery: New York, NY, USA, 1998; pp. 253–254. [CrossRef]
24. Dray, S.; Siegel, D. Remote Possibilities? International Usability Testing at a Distance. *Interactions* **2004**, *11*, 10–17. [CrossRef]

25. Bruun, A.; Gull, P.; Hofmeister, L.; Stage, J. Let Your Users Do the Testing: A Comparison of Three Remote Asynchronous Usability Testing Methods. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Boston, MA, USA, 4–9 April 2009; pp. 1619–1628. [\[CrossRef\]](#)
26. Radiah, R.; Mäkelä, V.; Prange, S.; Rodriguez, S.D.; Piening, R.; Zhou, Y.; Köhle, K.; Pfeuffer, K.; Abdelrahman, Y.; Hoppe, M.; et al. Remote VR Studies: A Framework for Running Virtual Reality Studies Remotely Via Participant-Owned HMDs. *arXiv* **2021**, arXiv:2102.11207.
27. Steed, A.; Archer, D.; Congdon, B.; Friston, S.; Swapp, D.; Thiel, F.J. Some Lessons Learned Running Virtual Reality Experiments Out of the Laboratory. *arXiv* **2021**, arXiv:2104.05359.
28. Burova, A.; Heinonen, H.; Becerril Palma, P.; Keskinen, T.; Hakulinen, J.; Opas, V.; Mäkelä, J.; Ronkainen, K.; Siltanen, S.; Raisamo, R.; et al. Towards Efficient Academia-Industry Collaboration Methodology: A Case Study of Joint Development of VR System. *Acad. Mindtrek* **2021**, in print.
29. McFadden, E.; Hager, D.R.; Elie, C.J.; Blackwell, J.M. Remote Usability Evaluation: Overview and Case Studies. *Int. J. Hum. Comput. Interact.* **2002**, *14*, 489–502. [\[CrossRef\]](#)
30. Andreasen, M.S.; Nielsen, H.V.; Schröder, S.O.; Stage, J. What Happened to Remote Usability Testing? An Empirical Study of Three Methods. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07), San Jose, CA, USA, 28 April–3 May 2007; Association for Computing Machinery: New York, NY, USA, 2007; pp. 1405–1414. [\[CrossRef\]](#)
31. Madathil, K.C.; Greenstein, J.S. Synchronous Remote Usability Testing: A New Approach Facilitated by Virtual Worlds. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11), Vancouver, BC, Canada, 7–12 May 2011; Association for Computing Machinery: New York, NY, USA, 2011; pp. 2225–2234. [\[CrossRef\]](#)
32. Madathil, K.C.; Greenstein, J.S. An Investigation of the Efficacy of Collaborative Virtual Reality Systems for Moderated Remote Usability Testing. *Appl. Ergon.* **2017**, *65*, 501–514. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Bolderston, A. Conducting a Research Interview. *J. Med. Imaging Radiat. Sci.* **2012**, *43*, 66–76. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Chapman, D.S.; Uggerslev, K.L.; Webster, J. Applicant Reactions to Face-to-Face and Technology-Mediated Interviews: A Field Investigation. *J. Appl. Psychol.* **2003**, *88*, 944–953. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Gruber, M.; Eberl, J.-M.; Lind, F.; Boomgaarden, H.G. Qualitative Interviews with Irregular Migrants in Times of COVID-19: Recourse to Remote Interview Techniques as a Possible Methodological Adjustment. *Forum Qual. Soz. Forum Qual. Soc. Res.* **2021**, *22*, 7. [\[CrossRef\]](#)
36. Weller, S. Using Internet Video Calls in Qualitative (Longitudinal) Interviews: Some Implications for Rapport. *Int. J. Soc. Res. Methodol.* **2017**, *20*, 613–625. [\[CrossRef\]](#)
37. Walsh, T.; Nurkka, P.; Walsh, R. Cultural Differences in Smartphone User Experience Evaluation. In Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia (MUM '10), Limassol, Cyprus, 1–3 December 2010; Association for Computing Machinery: New York, NY, USA, 2010; Volume 24, pp. 1–9. [\[CrossRef\]](#)
38. Gruenstein, A.; McGraw, I.; Badr, I. The WAMI Toolkit for Developing, Deploying, and Evaluating Web-Accessible Multimodal Interfaces. In Proceedings of the 10th International Conference on Multimodal Interfaces (ICMI '08), Chania Crete, Greece, 20–22 October 2008; Association for Computing Machinery: New York, NY, USA, 2008; pp. 141–148. [\[CrossRef\]](#)
39. Qian, J.; Chapin, A.; Papoutsaki, A.A.; Yang, F.; Nelissen, K.; Huang, J. Remotion: A Motion-Based Capture and Replay Platform of Mobile Device Interaction for Remote Usability Testing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2018**, *2*, 1–18. [\[CrossRef\]](#)
40. Fidas, C.; Katsanos, C.; Papachristos, E.; Tselios, N.; Avouris, N. Remote Usability Evaluation Methods and Tools: A Survey. In Proceedings of the 11th Panhellenic Conference on Informatics, Patras, Greece, 18–20 May 2007.
41. Kittur, A.; Chi, E.H.; Suh, B. “Crowdsourcing for Usability: Using Micro-Task Markets for Rapid Remote and Low-Cost User Measurements”, Conference on Human Factors in Computing Systems. 2008, pp. 453–456. Available online: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.532.2255&rep=rep1&type=pdf> (accessed on 2 May 2021).
42. Nebeling, M.; Speicher, M.; Norrie, M.C. Crowd Study: General Toolkit for Crowdsourced Evaluation of Web Interfaces. In Proceedings of the 5th ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS '13), London, UK, 24–27 June 2013; Association for Computing Machinery: New York, NY, USA, 2013; pp. 255–264. [\[CrossRef\]](#)
43. Henze, N.; Boll, S. Push the Study to the App Store: Evaluating off-Screen Visualizations for Maps in the Android Market. In Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '10), Lisbon, Portugal, 7–10 September 2010; Association for Computing Machinery: New York, NY, USA, 2010; pp. 373–374. [\[CrossRef\]](#)
44. Steed, A.; Ortega, F.R.; Williams, A.S.; Kruijff, E.; Stuerzlinger, W.; Batmaz, A.U.; Won, A.S.; Rosenberg, E.S.; Simeone, A.L.; Hayes, A. Evaluating immersive experiences during COVID-19 and beyond. *Interactions* **2020**, *27*, 62–67. [\[CrossRef\]](#)
45. Garcia Estrada, J.; Prasolova-Førland, E. Running an XR Lab in the Context of COVID-19 pandemic: Lessons Learned From a Norwegian University. *Educ. Inf. Technol.* **2021**, 1–17. [\[CrossRef\]](#)
46. Bhatia, S.; Björling, E.A.; Budhiraja, T. Exploring Web-Based VR for Participatory Robot Design. In Proceedings of the Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction (HRI '21 Companion), Boulder, CO, USA, 8–11 March 2021; Association for Computing Machinery: New York, NY, USA, 2021; pp. 109–112. [\[CrossRef\]](#)
47. Mathis, F.; Zhang, X.; O'Hagan, J.; Medeiros, D.; Saeghe, P.; McGill, M.; Brewster, S.; Khamis, M. Remote XR Studies: The Golden Future of HCI Research? In Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '21), Yokohama, Japan, 8–13 May 2021; ACM: New York, NY, USA, 2021.

-
48. De Ponti, R.; Marazzato, J.; Maresca, A.M.; Rovera, F.; Carcano, G.; Ferrario, M.M. Pre-graduation Medical Training Including Virtual Reality during COVID-19 Pandemic: A Report on Students' Perception. *BMC Med. Educ.* **2020**, *20*, 332. [CrossRef] [PubMed]
 49. Albastaki, A.; Hoggenmüller, M.; Robinson, F.A.; Hespanhol, L. Augmenting Remote Interviews through Virtual Experience Prototypes. In Proceedings of the 32nd Australian Conference on Human-Computer Interaction (OzCHI '20), Sydney, Australia, 2–4 December 2020; Association for Computing Machinery: New York, NY, USA, 2020; pp. 78–86. [CrossRef]
 50. Human Optimized XR Project. Available online: <https://humanoptimizedxr.org/wp/> (accessed on 15 August 2021).
 51. KEKO Smart Ecosystem Project. Available online: <https://kekoecosystem.com/> (accessed on 15 August 2021).
 52. Gabriel, A. Evaluating the Utility of Multi-User VR in Product Development. Master's Thesis, Aalto University, Espoo, Finland, 2020. p. 88. Available online: <http://urn.fi/URN:NBN:fi:aalto-202008234991> (accessed on 2 May 2021).
 53. Available online: <https://3dtalo.fi/designspace/> (accessed on 2 May 2021).
 54. Available online: <https://www.glue.work/> (accessed on 2 May 2021).
 55. Ries, E. *The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses*; Crown Business: New York, NY, USA, 2011.
 56. THL. Transmission and Protection—Coronavirus. Available online: <https://thl.fi/en/web/infectious-diseases-and-vaccinations/what-s-new/coronavirus-covid-19-latest-updates/transmission-and-protection-coronavirus/> (accessed on 2 May 2021).
 57. Thompson, S.M. Remote Observation Strategies for Usability Testing. *Inf. Technol. Libr.* **2003**, *22*, 22–31.