

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,000

Open access books available

148,000

International authors and editors

185M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Chapter

Usability Testing of Mixed Reality Scenarios: A Hands-on Report

*Robert Strohmaier, Gerhard Sprung, Alexander Nischelwitzer
and Sandra Schadenbauer*

Abstract

We would like to share our insights in designing, preparing, performing, and analyzing usability tests for multiple connected augmented reality and virtual reality applications as well as traditional mobile applications developed for a multimodal screening tool. This screening tool is under development at the University of Applied Sciences FH JOANNEUM in Graz, Austria. Several researchers from the departments of health studies and applied computer sciences are working closely together to establish a tool for early diagnosis of cognitive impairments to contribute to the management of dementia. The usability of this screening tool was evaluated by ten therapists paired with ten clients as testing group 1 and two usability experts in a separate test (group 2). In this chapter, we would like to describe why we use observed summative evaluation using the co-discovery method followed by post-task questionnaires for the first testing group. We are going to discuss the reasons for performing the cognitive walkthrough method as co-discovery with usability experts of testing group two as well. Furthermore, we describe how we use camera recordings (traditional cameras, 360-degree cameras), screen recording, and special tailor-made software to experience the screening process through the user's eyes.

Keywords: usability, user centered design, augmented reality, virtual reality, mixed reality, cognitive walkthrough, co-discovery, observation

1. Introduction

During this chapter we are going to discuss a usability study of a mixed reality scenario. Within the next pages, relevant literature is shown and insights into our ideas during planning of the usability study are given. As well as the subject of the test, the multimodal mixed reality scenario is going to be presented. In every section we tried to point out what is, in general, to do at a specific step or task during the usability processes. As well, we try to describe how we have done something (highlighted by the words "In our case") and, if applicable, we try to give suggestions, i.e., what can be improved (highlighted by the words "Potential for improvement").

First, we are going to clarify basic terms to establish a common understanding of the material at hand.

When it comes to classifying an application into the areas of augmented reality (AR), virtual reality (VR), and mixed reality (MR), Vogel et al. gives a good overview of how these techniques are seen in different disciplines [1].

1.1 Augmented reality (AR)

In our remarks, we refer to the definitions of Azuma [2] and the virtuality continuum of Milgram [3]. AR has to fulfill the following three points:

- **Virtual objects and virtual aid enrich the real environment**
The virtual objects and virtual aid have to be connected closely with the real environment. Virtual aid can be interpreted as visible content such as 3D models as well as invisible services like a digital tape measure to measure distances in the real environment.
- **Anchored in real space**
In our point of view, it is important that virtual objects and services are anchored to the real environment. This connection has to be stable and must not change during the use of the application.
- **Updates in real time**
All kinds of interactions should be computed in real-time. Furthermore, reactions to actions have to be rendered in real-time or, at least, with a minimal delay.

1.2 Virtual reality (VR)

Virtual reality can be seen as a completely virtual environment, where users are isolated from the real world and experience 3D models, spatial audio, and other digital content as well as other sensory stimuli. Ideally, the real world cannot be recognized at all. No real-world image, no real-world acoustic signal, no real-world smell, and no real-world tactile stimuli. Regarding the used hardware, this can be established completely or partly. Whereas using just a VR headset and moving within an environment like a therapist's treatment room, the immersion in the virtual environment is lesser. Regarding Milgram, we also tend to categorize our VR applications in the area of the virtual environment [4].

1.3 Mixed reality (MR)

Mixed reality is described as everything between the real and the virtual environment [3]. In addition, some major hardware and software companies tried to use this term to fit their hardware and software. In the following paragraphs, we are going to explain the elements of the conducted usability tests. All of them together are considered by us to be a mixed reality scenario.

1.4 Other important terms

Besides technical terms, during this chapter, we use many terms related to usability testing as well as terms related to the research project named SCOBES-AR (description of the project follows in the next section), itself. Therefore, within this

section, we are going to clarify the meanings of the most important terms. In the case of any confusion about the used terms later on, please refer to this section.

- Usability test
Structured process to assess how usable a computer-based system is.
- Usability method
Determines how a usability test must be carried out. E.g., heuristic evaluation or thinking aloud can be seen as a usability method.
- Usability study
If multiple usability tests are carried out to answer the same questions about the behavior of the users or the functionality of the product, we refer to this as a usability study.
- Expert (Usability expert)
A person who possesses profound knowledge about usability.
- Facilitator
A usability expert, who is in charge of planning, conducting, and analyzing usability tests.
- Observer
A person with at least basic knowledge about usability assists the facilitator during usability tests. They are mainly in charge of observing the tests and taking notes.
- Participant
A person who takes part in usability tests in the role of a user. This person must be recruited out of the target group(s) of the product to be tested. In our case, both therapists who are using the screening tool, as well as their clients being screened, can be participants in our usability study.
- Screening tool
A set of different screenings to, in our case, test physical and cognitive abilities in healthy persons.
- Screening
Activities to assess a person's abilities within a special scope (e.g. cognitive health, physical health).
- Task
One step in screening.
- Therapist
A health professional who offers some kind of health-related service. Therapist can utilize the screening tool developed by us.
- Client
A client takes advantage of a health-related service, provided by a therapist. They can be screened by a therapist.

1.5 The Project SCOBES-AR

Cognitive impairments like dementia are a tremendous challenge for affected persons, their relatives, and for a country's health system. To detect cognitive impairments at a very early stage might bring a lot of advantages for people and for administration. Therefore, we are working on a multimodal screening tool (MST) to detect cognitive impairments early. After identifying suitable screenings, we try to advance these screenings by digitalizing them [5].

To accomplish these goals, we have developed multiple digital applications. Some of them fit into the area of AR, some of them into VR and others are simple digital applications running on smartphones and tablet computers. All AR, VR, and traditional digital apps form the digital screening tool. We see this combination as a mixed reality scenario. This is the reason for using the term MR in the title of this chapter. More information about the constituent applications, whether they are AR, VR, or just digital, is given in the next section.

2. Elements of the usability tests

This chapter's main focus lies on usability and usability test methods. Nevertheless, a short overview of the single applications which together make up the screening tool is given within this section.

Overall, the screening tool consists of several screenings. At the beginning of each screening, anamnestic data is collected with the help of questionnaires presented on websites. Those questionnaires are not subject to the usability test. We tried to focus on the AR, VR, smartphone and tablet applications, and the process they are used (see **Figure 1**).

During the whole process, therapists have to take care of clients, they have to start the right software at the right time, prepare the proper hardware and, in case of malfunctions, they have to troubleshoot. Because of that, handling the hardware is also subject of the usability tests.

With the application running on the tablet computer, therapists have to log into the client using a quick response (QR) code. Then, they can start and stop the screenings described below. They can also visually monitor everything the client sees in the AR and VR headsets. They are able to monitor and annotate the screenings (see **Figure 2**).

After this short description of the application used by the therapists, we are going to introduce the screenings, which are the subject of the usability tests (see **Figure 1**).

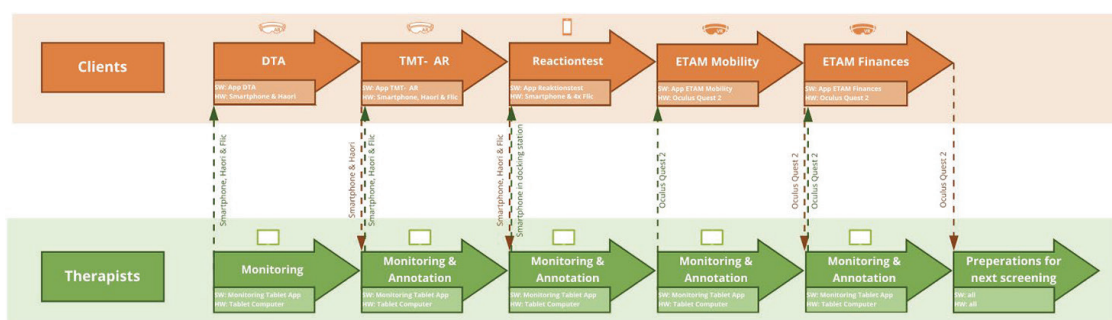


Figure 1.
Process of all screenings.



Figure 2.
Therapist during screening (monitoring and annotation).

2.1 Trail making test AR (TMT-AR)

Cognitive parameters are assessed during a process where clients have to select numbers and letters in ascending order [6, 7]. In the traditional form of the screening, these numbers are printed on a sheet of paper and pinpointed with a pencil, or the test is executed on a computer system.

In our case, numbers and letters are shown in virtual spheres, which are superimposed over the real world with the help of AR techniques. To accomplish this, the client wears a smartphone mounted in a Haori AR headset. To select single spheres, a flic2 button [8] is used. This can be seen in **Figure 3**.

2.2 Dual task assessment (DTA)

Gait parameters are measured by having the client walk a distance of 10 meters several times. If this screening were to be conducted traditionally, the distance would be marked on the floor. Therapists would use a stopwatch to measure elapsed time and would count the steps of the clients [9].



Figure 3.
Smartphone mounted in Haori headset worn by the client.

In our case, the client also wears a smartphone mounted in a Haori AR headset. A custom smartphone application utilizes AR techniques to measure distance, time, and footsteps.

2.3 Reaction test

During the traditional screening, reaction times of clients are measured by the talent diagnosis system (TDS) using the method “match 4 point” [10, 11].

In our case, a smartphone application is used to show visual stimuli. Depending on the stimulus, the client has to push the corresponding flic2 button. This application is just a smartphone app. No AR or VR techniques are used here.

2.4 ETAM mobility

The Erlangen test of activities of daily living in persons with mild dementia or mild cognitive impairment (ETAM) is carried out traditionally with printed pages. For the “mobility” element of this screening, clients are shown traffic scenes printed on paper. They have to argue about what people in these scenes are allowed to do [12].

In our case, clients wear an Oculus Quest 2 VR headset and find themselves within virtual environments, displayed as spherical 360-degree videos of traffic situations. Clients are able to look in all directions and have to interact with the system to decide where to go, when, and why (**Figure 4**).

2.5 ETAM finances

At this screening, the client’s abilities regarding recognition of prices and mental arithmetic are measured. Traditionally the client has to decide which products to buy using a very simple fake catalog of a grocery store (**Figure 5**). In addition, money for a shopping list has to be calculated exactly with real currency units [12].

In our case, clients accomplished these tasks in a virtual shop wearing the Oculus Quest 2 VR headset.

After introducing the project SCOBES-AR and the single screenings of the screening tool, we’d now like to give an overview of some promising usability methods to test its user-friendliness.



Figure 4.
ETAM mobility.

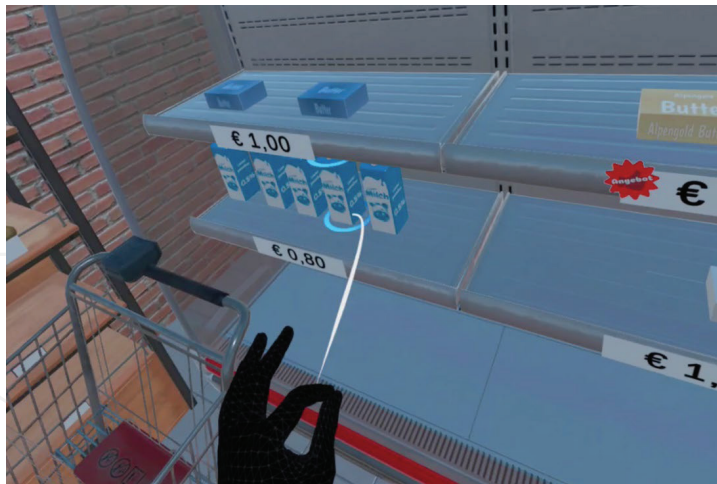


Figure 5.
ETAM finances - view from inside VR headset.

3. Selection of suitable usability testing methods for AR and VR

Based on usability literature and our experiences we have preselected several usability testing methods. Those will be discussed, especially regarding their use in AR, MR, and VR settings, within this section. After reading the following paragraphs it will be clear which methods we have chosen and why.

At the beginning of every usability study, the selection of suitable usability testing methods takes place. Usability testing methods mostly originated during the rise of desktop software and websites. Nowadays the amount of AR and VR applications increases steadily. In tandem, attempts to test their usability become increasingly common. A good point to start from is recent papers and articles, in which usability studies of AR applications [13, 14] and of VR applications [15, 16] are presented or compared. Here, it is essential to know how traditional usability tests work. Authors like Barnum [17] and Nielsen [18] as well as practical guides [19] should be considered, before applying or adapting a usability method.

In our point of view, feedback from real users is essential to develop a meaningful product. Additionally, the detailed feedback of experts is needed, when real users cannot imagine what could be improved and how this could be achieved. In this case, well-experienced usability experts are very valuable. Hence, we designed two usability studies. The first is with methods suitable for real users and the second just to get feedback from usability experts.

During the next paragraphs, we would like to give a short overview of promising usability testing methods for this need. Along with this overview, we focus on the differences of the traditional method and the usage for AR, VR or MR.

3.1 Heuristic evaluation

According to Nielsen, heuristic evaluation is an expert tool, wherein usability professionals analyze a product regarding a set of 10 heuristics. Those heuristics cover a large amount of usability factors. Analyzing a product with this 10-point list determines if the product is giving proper feedback to the users, is oriented along real-world metaphors, provides helpful error dialogs, and many other aspects [18, 20, 21].

Those 10 heuristics are suitable for traditional desktop applications and websites. But, when it comes to other domains, some of them might not be suitable for the task at hand. For this reason, different researchers tried to adapt the heuristics for their own use. When it comes to VR, Sutcliffe and Gault came up with 12 heuristics [22] whereas Rusu et al. tried to introduce 16 heuristics [23] followed by a 9-point catalog of Murtza, Monroe, and Youmans [24]. When analyzing AR applications, Gale et al. suggest a 26 points catalog [25] and Endsley et al. compare different approaches [26].

Some of the recently developed heuristics are comparable with some of the 10 heuristics developed by Nielsen. Showing the system status to users, giving appropriate feedback, and helping to recover from errors are some points that are still important nowadays. Additionally, the newer heuristics are influenced mainly by factors of the AR and VR hardware like comfort of wearing the hardware and the immersion into a realistic (AR) or virtual (VR) environment.

Conducting a heuristic evaluation in the domains of AR, VR, and MR is challenging. A key factor in this challenge is considering the product to be tested and which heuristics to choose or how to alter some existing catalog of heuristics.

In our case, we have analyzed existing heuristics for AR and VR products. However, those heuristics were not suitable for us. As the process of developing a custom set of heuristics seemed to be out of the scope of the project, we rejected the idea of using heuristic evaluation.

3.2 Co-discovery

Evaluating products with two persons at the same time might be a very good idea. If two people are needed for the product or service, it is a good idea, to also test the usability with two people. This method can also be combined with other methods [17].

Furthermore, if the AR, MR, or VR product has to be operated by two people, co-discovery can also be used in such a scenario [14].

In our case, the method of co-discovery is a perfect match. Since as the scenario of our project consists of two people, the therapist and the client, who are working closely together, we decided to use co-discovery. We were also in favor of combining co-discovery with another method. These thoughts are described in the following paragraphs.

3.3 Thinking aloud

Following the thinking-aloud method, users must fulfill several tasks using a product. The product can be a physical object like a coffee machine or desktop software, or a website. The participants of the usability test are encouraged to think aloud and verbalize their mental models while interacting with the product [17, 18]. This enables deep insights into the usability of the product. Usability experts observe the completion of the tasks and the thoughts of the users. For documentation, the whole process is often filmed. By inviting 3 to 5 users, the main usability issues can be found. While this method can be seen as a cheap method it is very unnatural for participants to think out loud [18].

When it comes to AR, MR, and VR settings, this method also seems to be very promising. Interacting with virtual objects in real or virtual worlds while verbalizing thoughts seems to be a good solution. For example, Boulder et al. used the thinking-aloud method when comparing usability tests of a car infotainment system in an MR scenario and in a real scenario [27]. Thinking aloud might work in a broad range of AR, MR, and VR scenarios. On the other hand, the limitations of the method have to

be considered. For example, if speech recognition is used to interact with the product, thinking aloud is inappropriate and may lead to accidental input. It must also be considered that thinking aloud produces a cognitive load on the side of the participants which can influence the results of a usability study.

In our case, it is crucial to get direct feedback from real users. Thinking aloud appeared to be a promising approach to accomplish this. But, as far as clients must carry out physical and especially cognitive screenings, we have seen huge problems in the thinking-aloud method. As well, usability tasks in addition to the tasks of the screening would be overwhelming for the clients and the therapists. Because of these reasons we did not apply this method to our usability studies.

3.4 Observation

The aim of observations is to study the users working with a product in a natural environment while not interrupting or disturbing them. Besides taking notes to gather qualitative feedback, videos can be recorded to analyze the observations. In some cases, the users can be interrupted to answer questions. According to Nielsen, three or more users have to be observed to get meaningful results [17, 18].

Because of the simple and product-independent setup, observations can also be conducted when evaluating AR [14, 28], MR, and VR products. Gaining qualitative feedback and impressions of the usage under normal conditions seems to be very promising.

In our case, observation fulfilled our requirements. It combines the view of real users doing everything which is needed for the real screenings without any influences of usability methods. Therefore, we applied this usability test method together with the co-discovery method. Therapist and client can perform a screening under realistic conditions while being observed to gather meaningful results regarding the usage of the different software applications and the process itself.

3.5 Questionnaires

Questionnaires can help to get structured feedback from participants before, during, and after a usability test. How many questionnaires to use and what they look like, is mostly defined by the product to be tested and the current stage of the product lifecycle. Questionnaires can be answered by the participants before (pretest), during (posttask) and after (posttest) a usability test. Whereas pretest questionnaires often aim to get more background information about potential users, posttask questionnaires relate directly to the completion of single tasks. Finally, posttest questionnaires are closely related to the goals of the full usability test [17, 19, 29].

When it comes to analyzing VR applications, VRUSE is an early attempt to provide a set of 100 questions structured in 10 parts [30]. Rather than answering the VRUSE questions within a Microsoft Excel spreadsheet, Putze et al. suggests implementing the questionnaires directly into the VR application. Here the questions are displayed within the virtual world, and they can be answered by using the VR controller, which is also used for the main application [16].

When using this approach in VR or AR persons with sight impairment must be considered. Besides, the implementation of a questionnaire into AR or VR software needs more resources as handling it like a traditional questionnaire.

Questionnaires can also be applied for AR applications. A good point to start research in this area is the overview provided by Pranoto et al. [14].

In our case, we have decided to use questionnaires after each usability observation to figure out if the perception of the usability test facilitator matches the self-perception of the participants. Considering the target groups and the duration of the whole screening process, we tried to come up with a simple and rather short questionnaire. We've analyzed the VRUSE questionnaire [30] and selected the following important aspects for the project SCOBES-AR:

1. Functionality
2. User input
3. Help and user manual
4. System output
5. Hardware
6. Immersion

According to these considerations we composed a set of questions. For each screening, the questions looked similar. This enabled us to draw conclusions between screenings. Because of the small sample of 10 therapists and 10 clients as participants, documentation was done in Microsoft Excel. This sample size is too small to draw conclusions from the questionnaires itself. However, in combination with the documentation of the observations, the questionnaires are able to help us to better understand what participants did and why they did it.

3.6 Cognitive walkthrough

When it comes to analyzing whether users are able to interact with a system in a proper manner, a cognitive walkthrough can be conducted. Inspecting especially the learnability of a system, action sequences are analyzed by usability experts. To be successful, the experts need information about the users of the system, and typical tasks for the system must be prepared. During the usability test, four main questions are being considered:

- Will the user try to get the right result?
- Will the user notice that the right action is available to them?
- Will the user associate the correct action with the expected result?
- When the right action is taken, will the user see progress being made toward the intended outcome?

Based on this analysis, suggestions were made on how to improve usability [17, 21, 31].

The main key is, whether using it for traditional digital products or AR or VR products, to make sure that the experts understand the users. This might be solved by personas and empathy maps, which can be created on the basis of focus groups [17–19, 31].

This method can just as well be used in AR, MR, and VR scenarios. The method may be altered slightly, but also can be utilized as described for other digital products [14, 15].

In our case, the expert view regarding those questions was exactly what we needed. To take into account the fact that at a real screening therapist and client work in pairs, we applied the cognitive walkthrough in form of a co-discovery.

After this discussion of usability methods, we are going into detail. Within the next section, the most important steps in preparing the usability studies are detailed. Please be aware, that more details can be found in usability literature.

4. Preparation phase

During this chapter, we explain the most important steps in designing a usability study. We will cover mainly the steps we have used for our study design and mention some other methods in short. Additionally, we will describe our lessons learned while designing the study for the mixed reality scenario.

4.1 Important documents

Usability experts suggest writing a guideline document for every usability test. Every consideration should be documented within this guideline. As well, all considerations during planning phase can be documented here.

Most times also a general data protection regulation (GDPR) [32] consent is needed where participants agree to the computation of their private data. Depending on the type of study and individual regulations in different countries, other documents like a data management plan and others might also be needed. Country depended regulations [33] must be met [17].

In our case: We've written an independent guideline for each usability study. Apart from goals, ideas, and to-do lists, the exact timetables of the usability tests were especially useful. Very helpful as well as contacting a law expert when it comes to documents relating GDPR. This step is strongly recommended.

4.2 What is the aim of the usability study?

The most important step in preparing the tests is to identify the aim(s) of the usability tests. Depending on the aim(s), different usability testing methods are suitable and different kinds of participants must be recruited. Also, the methods of documentation and presentation vary [17].

In our case: We wanted to see how clients and therapists handle the software and how they interact with each other. The typical usability of the software and the overall screening process had to be evaluated.

4.3 Recruitment of participants

Next should be considered whether the real users out of the target group(s) should be involved in the usability test, or if usability experts should analyze the product(s). Enough time for recruiting participants should be planned. Regarding the sample size, this step can take a long time. The number of participants strongly varies depending on the usability test method. For detailed information please have a look at the literature within chapter 3: Selection of Suitable Usability Testing Methods [17].

In our case: We wanted both the expert view and insights from the target groups. Therefore, we conducted two usability studies. The first study was conducted with 10 therapists and 10 clients, who worked together in pairs. The second study took place with two usability experts, one of them took over the role of the therapist and the other the role of the client.

It is good practice to inform participants during recruitment on the phone about what will happen during the usability tests (e.g., if they will be recorded with video cameras and microphones) and especially about GDPR. If this is not done in the worst case a participant shows up and does not agree to be recorded. In this case, the test ends before it begins, and the time of every person involved is wasted. Some of our participants stated at the beginning of the tests that they feel a little bit uncomfortable when being filmed. However, every participant agreed to sign the GDPR document. If they would not be informed beforehand, during the recruitment phase, some of them maybe would refuse to cooperate right at the beginning of the test.

4.4 Selection of the usability test personnel

The consideration about who should lead and observe the usability tests should be taken right at the beginning of the planning phase. Deep knowledge of usability testing is strongly needed especially for the facilitator. Ideally, the whole of the usability personnel contributes during the planning phase. As well it should be decided who is responsible for analyzing the findings after the usability tests [17].

In our case: We have decided to run the tests with just one person. Aside from other reasons, limited resources led us to this decision. This means the facilitator was in charge of all interactions with the participants, doing observations during the usability tests, and handling the hardware. Doing so can be described as possibly stressful for the facilitator. However, it is feasible. Simply ensure that this person is very well prepared and can handle hardware and malfunctions of the hardware in proper time by themselves.

Potential for improvement: If two or more participants work together on a usability test at the same time, each one should be observed by one distinct observer. The facilitator might take on the role of one observer in such a scenario. Activities like preparing documents or preparing the hardware and product(s) used in the usability test can be shared between those persons. In any case, observing the participants should be shared. Focusing on one participant while taking notes has a huge potential in increasing the quality of the notes whilst also reducing the time spent during analysis of the notes after.

4.5 Defining the usability test methods

The aim(s) of the usability study and the product to be tested has a huge impact on choosing the right testing method. Also, choosing more than one testing method can lead to improved results. Because of the variety of available testing methods and the novelty of testing methods for AR and VR this selection process is described in detail in chapter 3: Selection of Suitable Usability Testing Methods.

4.6 Defining the tasks

After deciding these details, the tasks of the usability studies and/or the questions of the questionnaires must be defined. Depending on the selected testing method(s) and the product to test, the tasks that the participants of the usability study have to

accomplish are very different. Always keep the aim(s) of the study in mind when it comes to creating the tasks.

Besides defining the tasks, it can also be necessary to prepare additional information. When using an expert usability method, it might be essential to present results of personas, empathy maps, or similar tests to them.

In our case: In usability study one, we wanted the therapists and clients to act as they do normally during such a screening. Therefore, no specific tasks were defined. The only task was to do the screening. Therapists were introduced to the screening tool in advance and most of them had also done several screenings before the usability test. The clients are instructed by the therapists and do not have to have special knowledge about the screening methods.

When it comes to the expert test, the whole story looks a little bit different. The usability experts do not have knowledge about the screening methods. Nor do they have domain knowledge in the field of the therapists like physical therapy, occupational therapy, dietetics, or speech therapy. To make them familiar with the screenings, the screening method itself and the therapeutic background were explained. Special functions regarding the software were not. As well they have to be provided with detailed information about the target groups in the form of personas and empathy maps.

Besides thinking about tasks themselves, how the task completion can be measured and evaluated has to be considered as well during the preparation phase. In our case, we gathered qualitative data during the observation and categorized it with a coding scheme, and prioritized them by importance and urgency (see 6.3 Notes of the Observations and Codes).

When questionnaires will be used, the questions should be developed during the planning phase. A scale to rate the answers must also be defined. As a scale we used a five-level Likert scale with the following items:

Strongly disagree - Disagree - Neither agree nor disagree - Agree - Strongly agree.

4.7 Defining the documentation method(s)

Last but not least, during preparation, the documentation and analysis of the usability studies have to be considered. It's important to get an overview of needed hardware and software. As well it is an important step in planning to decide, when and where to place which camera(s). This is going to be discussed in Section 6 Documentation and Analysis.

After writing about the preparation phase, we describe in the next section, how we have carried out the usability tests.

5. Carrying out the usability tests

Below we try to give an overview of steps, which were very important for our usability studies. Again, like in the previous sections, more detailed information can be found in usability literature written by Barnum [17] or Nielsen [18]. The focus of this chapter stays on a hands-on report, where we try to show what we have done and why.

5.1 Usability study 1: observed summative evaluation

The first usability study was done in the form of an observation followed by a posttest questionnaire. It was combined with screenings under real conditions.

10 therapists and 10 clients were recruited for the usability tests. They had been informed about GDPR in advance and were willing to be filmed when they arrived. Starting with an introduction to usability and the reasons for the usability study, the facilitator activated all the video cameras and screen recorders. By clapping his hands, the audio signal for synchronization was given. Retreated to the background, the facilitator started taking notes. From screening to screening the facilitator had to change the positions of the cameras. The rest of the time, he tried to be in the background not influencing the screening.

After the screenings, a posttest questionnaire was handed out, where therapists and clients got different questions. The aim of the questionnaires was to gather additional feedback and to better understand the actions taken by the users during the usability test.

Potential for improvement: It might be better if every person who is to be observed would get a distinct observer. Observing two persons at a time is possible. But more detailed notes can be taken, if one observer can concentrate on a single user.

5.2 Usability study 2: cognitive walkthrough as a Co-discovery

As mentioned before, cognitive walkthrough was chosen as a method to gain deep insights into the usability from experts. Because of the fact that during the screenings therapists work in pairs with their clients, the method of co-discovery was also utilized. So, two usability experts were invited to analyze the product using the method of a cognitive walkthrough working in pairs, which is typical for the co-discovery method.

The agenda of the cognitive walkthrough was quite tough. All of the four screenings which utilize AR and VR have to be tested by experts. The experts themselves have profound knowledge about usability, computer science, interaction design, and didactics. But what they do not have is knowledge about medicine, health care, and traditional screening methods. Therefore, they must be introduced to every single screening, while not showing them too much detail about the software before the usability test starts. In doing this, the facilitator followed this outline, whereas 0x is a placeholder for all 4 screenings. So those three parts are carried out 4 times in the test.

- Introduction to the project
- Introduction to the personas and empathy maps
- Introduction to screening 0x
- Carrying out screening 0x
- Discussion about the main issues during screening 0x
- Closing remarks

While the screening took place, everything was filmed by a screen recorder with an activated microphone in the therapist's tablet computer. Notes were taken by the facilitator during the screenings as well as during the discussion of each screening. During the phase of discussing a screening, the screen recording was very valuable. Uncertainties during the discussion about what was happening before the screening could be clarified easily.

A very important point to discuss when planning the agenda are the breaks. The whole usability test session took us about 6 hours. It is really hard for the experts and for the facilitator to stay focused. A coffee break, some sweets and a lunch with time for networking were very important to stay on task.

More detailed information about the process of the cognitive walkthrough itself can be found in the next section, where one of the usability experts outline his experiences during the test.

Potential for improvement: At the beginning, special terms which are used within the product should be presented and clarified. A small, printed dictionary might help a lot, too, when wording is not clear.

It turned out, that during expert tests the experts also tend to give feedback about the usability test itself. One of their suggestions was given in relation to the empathy maps. Those should be extended to mood boards, where personas can be seen in different situations which are important for the product. Grouping and categorizing feelings take a very important place as well and should be designed in a proper way.

The section below is written by one of the participating usability experts and should give insights into the perspective of the experts during the usability test.

5.3 Impressions from participating usability experts

The task was to review the usability of the dashboard application used by therapists and the applications used by the clients during a screening. Furthermore, the communication and interaction between therapists and clients have to be assessed.

Right at the beginning of the usability test, it became clear that a lot of used terms were not clear at all. Wording is a very important point in communication.

We have to take into account a large number of different roles and needs, wishes, and fears associated with them. The roles that we two usability experts are supposed to assume are extremely differentiated. The introduction to the personas was essential to conducting the cognitive walkthrough. The phase of discussing persona descriptions and empathy maps took much longer than planned. The facilitator accepted these circumstances and altered the complete test schedule on the fly. Without deep knowledge about the target groups, the whole test might be completely senseless. Accordingly, it is a very unusual constellation in which a conventional setup would not have the desired success.

After studying and discussing personas and empathy maps in detail, we decided to change roles after each screening. This decision was made because our experiences in AR and VR methods were seen as helpful for the test settings. Also, more precise and differentiated findings were expected. This alternation was very helpful to act and discuss on a meta-level. Therefore, discussions not only about one single screening were held. Moreover, the complete screening tool was present during all steps of the usability test and was evaluated and accessed as a whole product.

6. Documentation and analysis

When it comes to taking notes using a coding scheme, recording several videos of the observation, synchronizing all the media created during the usability tests, and documenting and presenting the gathered and aggregated data, special usability testing software can be very handy. Unfortunately, when using usability software, costs emerge most times. If the budget for usability tests is severely limited, and you

are familiar with spreadsheet software like Microsoft Excel or something similar, you can use that software to design customized observation sheets, which is shown later. In the following section, the process of taking videos for documentation and analysis purposes will be discussed.

6.1 Interactions in the real world

Depending on the format of documentation, different pieces of hardware, like video cameras, microphones, and software (like special screen recorders) are needed. If interactions in a real environment are important, digital video cameras can be used. If interactions are distributed around in the real environment, a camera that can record spherical 360-degree video [34–36] of a whole scene at once might prove to be very useful [18].

In our case, we were interested in task completion in the smartphone, tablet, AR, and VR applications. It was very important for us to see the connection between tasks in the software and interaction in the real environment. Therefore, interactions in the real environment were recorded by two cameras facing each other to cover the whole test area. Additionally, we recorded spherical 360-degree video with the Ricoh Theta Z1 [37]. For each screening, the setting, shown as an example in **Figure 6**, has to be adapted a little bit.

But, recording video in the real world is not enough, when it comes to testing the usability of AR and VR scenarios.

6.2 Seeing through the users eyes

To see what users see is essential to understand interactions of users acting in AR and VR Applications. In the area of VR applications, screen recording software is sufficient most of the time. Hardware like the Oculus Quest 2 allows screen recording directly in the menu of the headset [38].

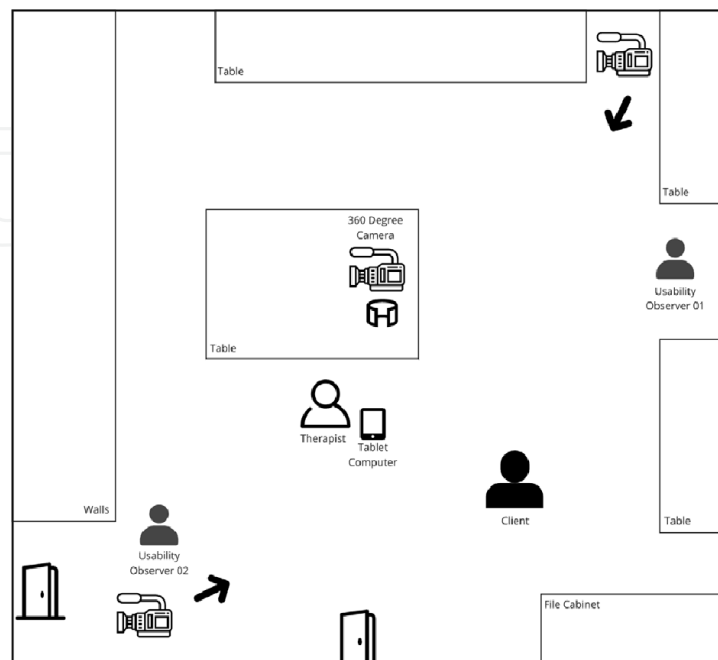


Figure 6.
Positions of the video hardware for the VR screenings.

When thinking about AR scenarios, screen recorders reach their limits. Depending on the techniques used to superimpose the real world, screen recorders might not be able to be used. When working with optical see-through devices like the Microsoft HoloLens a technique called spectator view comes into account. This technique allows us to record not only the virtual objects rendered into the sight of users but rather the whole virtual and real environment [39].

In our case, when it comes to interaction with the software, we have used Android's built-in screen recording tool on the therapist tablet computer. This solved two problems at once: within this screen recording we can reconstruct when which button was pressed by the therapist. As well, due to the fact that the therapist sees inside the tablet software exactly what the client sees, the view through the user's eyes is also recorded without any further screen recording software.

After the usability tests, the video material of all cameras must be synchronized and put together in a single video (**Figure 7**). A situation where this is not possible is with spherical 360-degree videos. They need special players to be shown most times. Synchronization of traditional recorded videos and screen videos works well with simple audio signals like clapping hands several times. Therefore, the microphones for each video camera and the ones connected to screen recorders must be activated. The step of synchronization should be done, to have one single video stream to have a look at, when the notes, which are going to be described in the next section, show interesting behavior.

6.3 Notes and codes

Before taking notes, a coding scheme should be created. As a coding scheme, we usually see a short description (1 or 2 words) or title to help to categorize all possible different findings during the usability observation. It can be very helpful to discuss

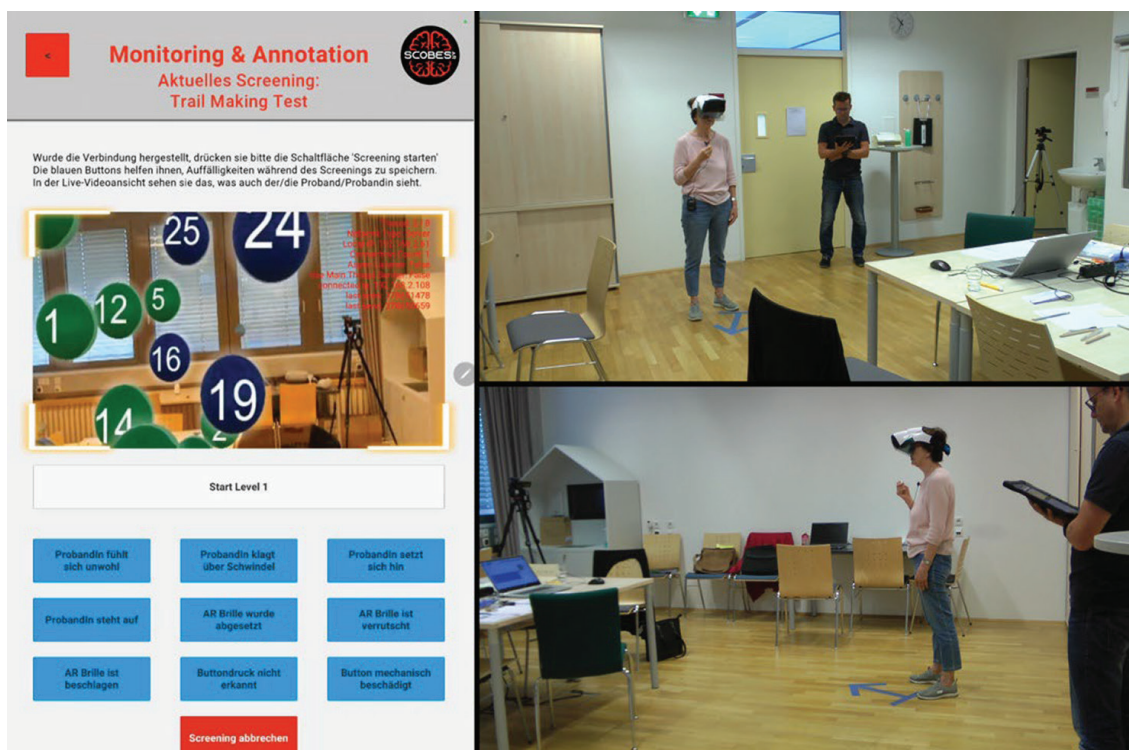


Figure 7.
All video-streams combined.

Updates on Software Usability

The screenshot shows an Excel spreadsheet with the following data:

Time	Timecode	Type (Code)	Who	Important	Urgent	Description
10:32:54	10:32:54	START				
10:35:40	00:02:46	Start DTA				
10:40:43	00:07:49	Function	Therapist	5	5	
10:42:02	00:09:08	Interaction	Client	1	3	
10:46:07	00:13:13	Ende DTA				
10:49:08	00:16:14	Start TMT				
10:51:42	00:18:48	Configuration	Therapist	4	1	
10:53:21	00:20:27	Tutorial	Both	2	1	
10:53:54	00:21:00	Interaction	Client	2	5	
10:54:17	00:21:23	Interaction	Client	1	3	

Figure 8.
Suggested use of excel spreadsheet for observations.

The left screenshot shows the formula in cell A9: `=IF(C9<>"",IF(A9<>"",A9;NOW());"")`. Red arrows and numbers 1, 2, and 3 point to the formula bar, the formula itself, and the 'NOW()' function respectively.

The right screenshot shows the formula in cell B10: `=A10-B9`. Red arrows and numbers 3 and 4 point to the formula bar and the formula itself respectively.

Figure 9.
Automatic creation of timestamps.

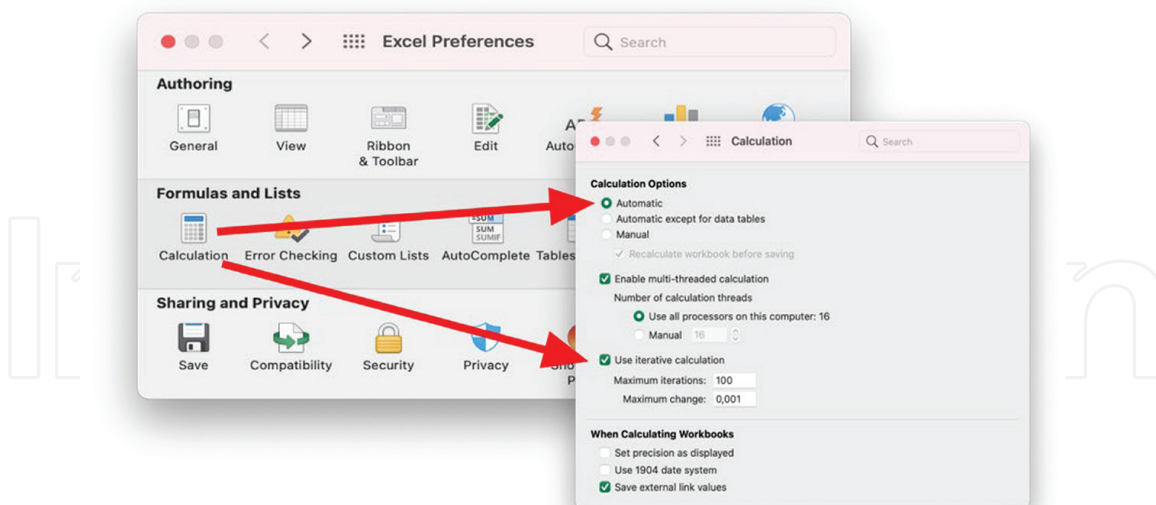


Figure 10.
Calculation options.

these codes with the development and usability team. Here, a pre-test is essential, to test and alter the coding scheme if necessary. Most times it is necessary.

The structure illustrated in **Figure 8** was very helpful for us. In column A the time of day was entered automatically when the code of the observed item was entered in the cell in column C. As well, based on the start time of the observation (cell A9) the corresponding timecode was calculated and entered automatically in the corresponding cell in column B. How to accomplish these two calculations in Microsoft Excel (Version 16.62 for Macintosh computers) is shown in **Figure 9**. The Timecode is extremely important when it comes to analyzing the video recordings. As mentioned in the section about the documentation using video an acoustic signal can be used to synchronize the recordings. At the moment of this signal, the first code (code START) has to be entered in column C9. From now on, the timecode of the observations is exactly the same as the timecode of the documentation of the video recordings.

If the calculations are not executed immediately after entering the code in column C, open Excel preferences and set in the menu calculation the calculation option to automatic. The option “Use iterative calculation” should be activated too. Otherwise, a warning about circular references might prohibit entering values. The appropriate setting for both options is illustrated in **Figure 10**.

7. Conclusions

Analyzing usability is a challenging task, whether the item to test is a traditional desktop application or a hardware-intensive MR scenario. Traditional usability methods, like the ones presented in this chapter, are useful and established. So, also for new purposes like AR, VR, and MR the traditional methodologies can be very helpful, at least as a starting point for research on new methods or to alter traditional methods to fit the items to test.

Furthermore, seeing through the user’s eyes has a tremendous impact on usability studies, when it comes to testing usability of AR, VR, and MR applications and scenarios. We were sure that it is crucial for the whole project, that therapists can exactly

see what clients see. But, also for the analysis of the usability studies, it was extremely helpful to watch the screen recording of the tablet software, where the viewport of the clients was shown. It assisted us in discussing usability issues with the experts after each screening.

In addition to that, we must emphasize, how important it is to inform participants during recruiting about the procedure of the usability test and, if applicable, that they will be filmed.

Regarding the process of observing a co-discovery, we can strongly recommend recruiting at least one observer per participant. The facilitator can take both roles, of course, ensuring that one observer is available per participant. The needs of the participants can, in this manner, be handled easier and observation results may be more detailed.

Last but not least, never underestimate the importance of detailed and lively persona descriptions. Combined with empathy maps they are a great tool, to help someone to take over the role of users from the target group(s). And, as the usability experts stated, empathy maps and personas can be enriched by pictures in the style of a mood board.

Acknowledgements

We'd like to thank all of our colleagues at the University of Applied Sciences FH JOANNEUM, who worked enthusiastically on the project SCOBES-AR and enabled the creation of the screening tool as well as the development of the software and the usability studies. Beginning with the project manager Wolfgang Staubmann, all other colleagues are presented in alphabetical order: Hannes Aftenberger, Konrad Baumann, Monica Christova, Theresa Draxler, Brigitte Loder-Fink, Bernhard Guggenberger, Ulrike Hauser, Nina Maas, Bianca Fuchs-Neuhold, Christoph Palli, Rene Pilz, Lucia Ransmayr, Helmut Simi, Anna Steiner, and Elisabeth Url.

As well we'd like to thank our students Miriam Grainer, Stefan Krasser, Lukas Bichler, Gerhard Zwitkovits and Andreas Krejan, who developed very helpful prototypes as starting points for the screening tool.

For proofreading and important remarks, we'd kindly like to thank Skye Sprung.

Finally, we thank the Austrian Institute of Technology for providing the backend system and securely hosting the collected data.

The project SCOBES-AR is funded by the Austrian Research Promotion Agency (FFG) under grant number 866873.

IntechOpen


IntechOpen

Author details

Robert Strohmaier*, Gerhard Sprung, Alexander Nischelwitzer
and Sandra Schadenbauer
FH JOANNEUM, Institute of Business Informatics and Data Science, Graz, Austria

*Address all correspondence to: robert.strohmaier@fh-joaanneum.at

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Vogel J, Koßmann C, Schuir J, Kleine N, Sievering J. Virtual- und augmented-reality-definitionen im interdisziplinären Vergleich. In: Thomas O, Ickerott I, editors. Smart Glasses. Berlin, Heidelberg: Springer Berlin Heidelberg; 2020. pp. 19-50
- [2] Azuma R. A survey of augmented reality. Presence: Teleoperators and Virtual Environments. 1997;**6**(4):355-385
- [3] Milgram P, Takemura H, Utsumi A, Kishino F. Augmented reality: A class of displays on the reality-virtuality continuum. In: 1995 SPIE Proceedings Vol. 2351: Telemanipulator and Telepresence Technologies. Boston, MA. 1995. pp. 282-292
- [4] Jerald J. The VR Book: Human-Centered Design for Virtual Reality. New York, San Rafael, California: Association for Computing Machinery Morgan & Claypool Publishers; 2016
- [5] FH Joanneum Gesellschaft mbH. Mixed Reality Prototype of Multimodal Screening for Early Detection of Cognitive Impairments in Elderly: Protocol Development and Usability Study [Internet]. clinicaltrials.gov; 2022. Report No.: NCT05403814. Available from: <https://clinicaltrials.gov/ct2/show/NCT05403814>
- [6] Tischler L, Petermann F. Trail making test (TMT). Zeitschrift für Psychiatrie, Psychologie und Psychotherapie. 2010;**58**(1):79-81
- [7] Rasmusson DX, Zonderman AB, Kawas C, Resnick SM. Effects of age and dementia on the trail making test. The Clinical Neuropsychologist. 1998;**12**(2): 169-178
- [8] Flic 2|The Smart Button for Lights, Music, Smart Home and More. Flic Smart Button. 2022. Available from: <https://flic.io/>
- [9] Hunter SW, Divine A, Frengopoulos C, Montero OM. A framework for secondary cognitive and motor tasks in dual-task gait testing in people with mild cognitive impairment. BMC Geriatrics. 2018;**18**(1):202
- [10] TDS Hardware 2011. 2022. Available from: <http://www.werthner.at/tds/tds-german/tds-hardware1.htm>
- [11] TDS Test Stand 2011-11. 2022. Available from: <http://www.werthner.at/tds/tds-german/tds-test1.htm>
- [12] Luttenberger K, Reppermund S, Schmiedeberg-Sohn A, Book S, Graessel E. Validation of the Erlangen test of activities of daily living in persons with mild dementia or mild cognitive impairment (ETAM). BMC Geriatrics. 2016;**16**(1):111
- [13] Merino L, Schwarzl M, Kraus M, Sedlmair M, Schmalstieg D, Weiskopf D. Evaluating Mixed and Augmented Reality: A Systematic Literature Review (2009-2019). In: 2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR). Porto de Galinhas, Brazil: IEEE; 2020. pp. 438-451
- [14] Pranoto H, Tho C, Warnars HLHS, Abdurachman E, Gaol FL, Soewito B. Usability Testing Method in Augmented Reality Application. In: 2017 International Conference on Information Management and Technology (ICIMTech). Yogyakarta: IEEE; 2017. pp. 181-186
- [15] Karre SA, Mathur N, Reddy YR. Usability Evaluation of VR Products in Industry: A Systematic Literature Review.

In: Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing. Limassol Cyprus: ACM; 2019. pp. 1845-1851

[16] Putze S, Alexandrovsky D, Putze F, Höffner S, Smeddinck JD, Malaka R. Breaking the Experience: Effects of Questionnaires in VR User Studies. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. Honolulu, HI, USA: ACM; 2020. pp. 1-15

[17] Barnum CM, editor. Usability Testing Essentials. 2nd ed. Morgan Kaufmann; 2021

[18] Nielsen J. Usability Engineering. San Diego, CA: Academic Press; 1993

[19] Richter M, Flückiger MD. Usability und UX kompakt. Berlin, Heidelberg: Springer Berlin Heidelberg; 2016

[20] Experience WL in RBU. 10 Usability Heuristics for User Interface Design. Nielsen Norman Group. 2022. Available from: <https://www.nngroup.com/articles/ten-usability-heuristics/>

[21] Sarodnick F, Brau H. Methoden der Usability Evaluation: wissenschaftliche Grundlagen und praktische Anwendung. Bern: Huber; 2006. p. 251

[22] Sutcliffe A, Gault B. Heuristic evaluation of virtual reality applications. *Interacting with Computers*. 2004;**16**(4):831-849

[23] Rusu C, Munoz R, Roncagliolo S, Rudloff S, Rusu V, Figueroa A. Usability heuristics for virtual worlds. In: AFIN International Conference on Advances in Future Internet. France: Nice/Saint Laurent du Var; 2011

[24] Murtza R, Monroe S, Youmans RJ. Heuristic evaluation for virtual reality

systems. Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 2017;**61**(1):2067-2071

[25] Gale N, Mirza-Babaei P, Pedersen I. Heuristic Guidelines for Playful Wearable Augmented Reality Applications. In: Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play. London United Kingdom: ACM; 2015. pp. 529-534

[26] Endsley TC, Sprehn KA, Brill RM, Ryan KJ, Vincent EC, Martin JM. Augmented reality design heuristics: Designing for dynamic interactions. Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 2017;**61**(1):2100-2104

[27] Bolder A, Grünvogel SM, Angelescu E. Comparison of the usability of a car infotainment system in a mixed reality environment and in a real car. In: Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology. Tokyo Japan: ACM; 2018. pp. 1-10

[28] Cavalcanti VC, de Santana MI, Gama AEF, Correia WFM. Usability assessments for augmented reality motor rehabilitation solutions: A systematic review. *International Journal of Computer Games Technology*. 2018;**2018**:1-18

[29] Geisen E, Bergstrom JR. Usability Testing for Survey Research. Cambridge, MA: Morgan Kaufmann Publisher; 2017

[30] Kalawsky RS. VRUSE—A computerised diagnostic tool: For usability evaluation of virtual/synthetic environment systems. *Applied Ergonomics*. 1999;**30**(1):11-25

[31] Barnum CM. Usability Testing and Research. New York, N.Y: Longman; 2002

[32] Pabst M. General Data Protection Regulation (GDPR) – Official Legal Text [Internet]. General Data Protection Regulation (GDPR). 2022. Available from: <https://gdpr-info.eu/>

[33] Unternehmensberatung A. Datenschutz-Grundverordnung (DSGVO) - JUSLINE Österreich [Internet]. 2022. Available from: <https://www.jusline.at/gesetz/dsgvo>

[34] What is Samsung Gear 360 camera?. Samsung uk. 2022. Available from: <https://www.samsung.com/uk/support/mobile-devices/what-is-samsung-gear-360-camera/>

[35] MAX 6K Waterproof 360-Degree Action Camera | GoPro. 2022. Available from: <https://gopro.com/en/us/shop/cameras/max/CHDHZ-202-master.html>

[36] 360-degree camera RICOH THETA. 2022. Available from: <https://theta360.com/en/>

[37] RICOH THETA Z1. 2022. Available from: <https://theta360.com/de/about/theta/z1.html>

[38] Oculus Quest 2. Oculus Quest 2: Unser bisher bestes, neues all-in-one VR-Headset | Oculus [Internet]. Oculus Quest 2. 2022. Available from: https://www.oculus.com/quest-2/?locale=de_DE

[39] Blog MD. How-to: Spectator View, a new tool to help others see what you see in HoloLens [Internet]. Microsoft Devices Blog. 2017 2022. Available from: <https://blogs.windows.com/devices/2017/02/13/spectator-view-new-tool-help-others-see-see-hololens/>