UIWatch: An immersive virtual smartwatch VR menu interaction study

Thomas Jalving 6056075 (Universiteit Utrecht)

2023

1 Introduction

Virtual Reality (VR) headsets have been around since the 1960s, with Ivan Sutherland creating one of the first Head Mounted Display (HMD) systems as shown in Figure 1. Recent developments have increased the usability of the VR platform, with devices such as the Oculus Quest (Figure 2), released in 2019, propelling the capabilities of VR forward. VR can now be used time and place independent on standalone headsets that do not require a tethered connection to a computer. With these advancements, the price of VR systems has also significantly dropped making them more accessible to a broader audience. These advancements sparked interest in research into what VR can do, with education and business being the biggest areas of interest.

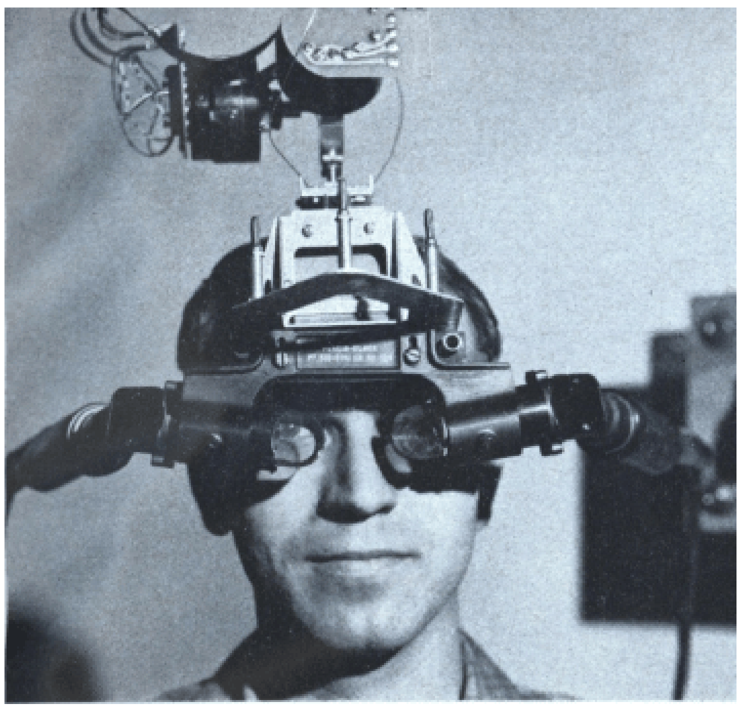

Figure 1: A photo of someone using the Sword of Damocles HMD which was created by Ivan Sutherland.

VR is being used more and more for educational applications in the past few years [47]. This growing interest demands more fundamental knowledge of best practices in educational VR. Literature on educational VR applications shows that an increased amount of immersion has a positive influence on learning outcomes [29]. But this immersion is broken when text and numbers clutter the 3D scene [9]. Researchers [40] also concluded that User Interface (UI) and User Experience (UX) have a direct influence on learning performance in VR. So it

Figure 2: A render of the Oculus Quest created by the company Meta.

seems that it is of importance that the UI is immersive and designed well.

In the past, many VR interaction methods have been proposed for UI. These methods range from simple raycast menus adapted from 2D [33] to multi-level depth based radial menus [17]. Those researchers created a foundation for UI and interaction design in VR, with a focus on usability and user preference. To date, almost no literature exists that combines both immersion and UI design. Most studies place focus on one or the other.

Since there is no already existing researched immersive interaction method, previous research into how to achieve immersion in general needs to be used to design one. Thus in this study I propose a new interaction method: UIWatch. A virtual smart watch that can be used to interact with the UI in an immersive way by using theory on immersion, which will be explained in more detail in Section 3. A method that might be useful for future educational VR applications and contributes to a better understanding of immersive UI elements. The goal of this study is to evaluate this new method, test different variants and compare it to the most common method: raycasting. The evaluation will be done on both immersion and usability.

This UIWatch is tested at the Dutch Police Acedemy (PA). VR is a new tool in their educational pipeline thus they are interested in doing research into best practices. They would like to make things such as gun safety training more accessible and cost effective. They would also like those types of educational materials to be able to be used at the pace of the student. Allowing some students to do more training on the things they find difficult.

2 Literature Research

This section will first contain a brief literature overview to prove that immersion is important in educational VR applications. Then it will display literature on how to achieve higher levels of immersion using the human senses in VR applications to be able to explain how the design of UIWatch could work in a later section. Lastly, an overview will be shown of recent research on the best practices of UI design in VR and interaction with them.

The literature research was performed by using Google Scholar. The following starting search terms were used after which other search terms followed where necessary:

1. "Virtual Reality" Interaction
2. "Virtual Reality" Menu Interaction
3. "Virtual Reality" UI
4. "Virtual Reality" Immersion
5. "Virtual Reality" Immersion Education

Quotation marks were used to force Google to search for Virtual Reality as a single word. Papers were filtered on date first, then title then the abstract to determine relevancy. Only papers on immersive Virtual Reality were considered. Relevant papers were read fully.

2.1 Terminology

It is important to define the term immersion since it is often mistakenly used interchangeably with the term presence. Immersion in this case refers to the user's engagement with the VR system that results in being in a flow state. Immersion is positively affected by things such as a higher display resolution, better graphics or better haptics [39]. Presence is defined as the sense of being in the virtual world [6] and how it is perceived [38]. Using these definitions, menu interactions can be linked to immersion as they are not meant to entice emotion. Menu interactions are also not open for a wide range of different perceptions.

Menu interaction will be seen as a subset of UI that only deals with menus.

2.2 Immersion in educational VR

VR in education as a topic took off in the early 1990s [10] [11], even catching the attention of the US military organisation DARPA [49] later on in the same decade. Although the research back then did not use the term immersion, it did conclude that VR is an interesting new educational platform that places the learning closer to the learning material.

Earlier VR research papers were looking at applying VR in educational fields[32]. Citing the capability of VR to bring the learner closer to the material and the provided high-level interaction were reasons why people enjoyed using it. And researchers [27] also started to look on the psychological side of using virtual environments.

With the launch of the first affordable devkit HMD in 2013, the Oculus Rift DK1, VR became more accessible than ever. A 2015 survey paper [18] mentions an increase in the amount of research papers thanks to the release of this headset. An interesting fact is that the research they found mainly focused on vocational training aimed at situations that are dangerous or hard to access. The research on those training applications are looking at immersive experiences as they try the simulate the dangerous situations as best as possible. This paper also notes that most of the educational research has been done on adult training.

A different survey paper from 2018 [29] found that in general the conclusion was that more immersion had a positive influence on learning outcomes. More immersive experiences were also taken more seriously by the users and VR was found to be especially useful in training situations that were stressful and required a good understanding of a 3D environment. For example, when comparing 360° video to the same video on a 2D screen, students were not only more motivated to learn about new topics in the VR environment but also showed better learning outcomes [36]. The positive effect of immersion on learning outcomes and learning performance has been found by a lot of researchers in the past decades across various learning fields [16] [42] [14] [13]. Overall the research strongly supports the use of VR in training simulations where immersion and realism are of importance.

A survey paper from 2020 [35] states that the research interest in using educational VR has never been bigger. However, the majority of the articles they found indicated that it is still in its experimental stages. An interesting fact is that most research is exploratory with engineering and computer science having the most papers. As with the research from 2013 to 2019, papers from 2019 onward strongly support the use of VR in education, especially with realistic and immersive simulations. Overall research from the past decades shows the importance of immersion in educational VR applications and a direct link to how well students learn from VR experiences.

2.3 How to achieve immersion

This section will contain literature on the different senses used in VR (visual, auditory and haptic) and their influence on the level of immersion a user is experiencing. The level of experienced immersion can be influenced by other factors such as the narrative [23] but these factors are outside the scope of this research.

A couple of different things are needed to achieve a higher level of immersion in VR applications. First of all, the visual fidelity of the headset is of importance. The more realistic the graphics, the higher the level of immersion. Researchers [9] state that things like the field of view, display resolution and realism of lighting have a direct influence on visual immersion. But visual immersion is only one part of a complete picture. Note that presence and immersion have been used to describe different things throughout different literature and that

I include UI with immersion as it is an important part of how an application feels and works.

Accompanying visual immersion is auditory immersion. Audio does not seem to have a large impact on the level of immersion a user experiences in immersive VR applications [7]. However, a complete absence of audio does decrease immersion and can even increase negative things like motion sickness [37]. Making the audio spatially localized, in other words 3D audio, does seem to add immersion and task completion competence [46]. In general, audio is a very small part of immersion but it still needs to be addressed in an application.

Another aspect to immersion is the sense of touch, also called haptics. Most of the research into this topic has tested different kinds of controllers or wearables to simulate certain forces or objects. An experiment showed that active haptics, where the controllers and wearables apply actual forces such as vibration, do not improve immersion significantly more than passive haptics [44]. Instead matching the weight and inertia of the object you are trying to simulate is more important for immersion. VR controllers do not need to match their virtual counterparts perfectly [20]. Changing the weight distribution also changes the users perception of what shape they are holding, which was proven by for example the Shifty experiment [50]. Haptics can also be faked by the use of visual cues that simulate physical forces[8]. Buying expensive special haptic controllers is not feasible for most use cases but haptics do offer a lot of benefit. A possible solution to this problem could be to use the users own body to generate haptic feedback since no extra hardware components are required for that.

Haptics seem to have more impact on presence than visuals but using multimodal feedback is the best. The more stimuli and the more sophisticated the feedback is, the higher the amount of experienced presence [22]. To achieve higher presence it is thus needed to stimulate as many of the human senses as possible.

2.4 VR menu interaction

Interaction is one of the main areas of interest for research into VR. With each new generation of headsets, new hardware capabilities are unlocked. Field of view and the display resolution increase. New types of controllers, and even hand tracking got added to the newest headsets. Thus each new generation sparks more interest into looking at new interaction types. One type of VR interaction is that with graphical elements displaying information, or menu interaction. Think of interactive buttons and text floating in mid air in the virtual environment.

Research on menu design in VR ranged from simple 2D raycast menus such as the one from Monteiro et al. [33] to much more complex multi-level depth based radial menus [17]. These two examples can be seen in Figure 3 and Figure 4. In general, these papers translated 2D concepts into a flat panel floating in the VR space.

While a fully 2D interface seems to be the most intuitive and easy to work

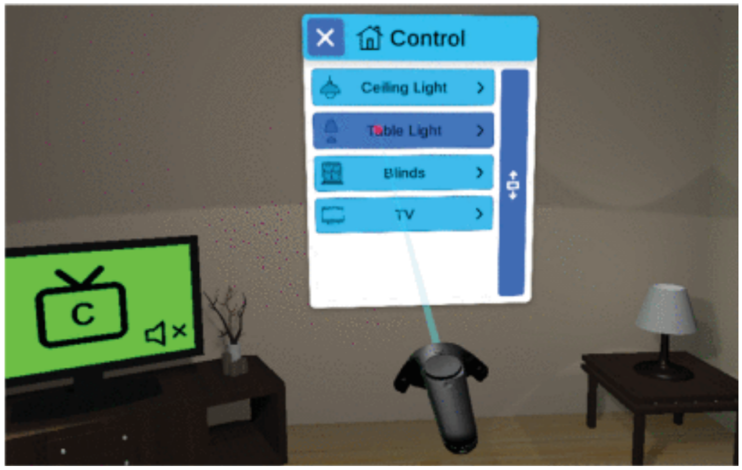

Figure 3: Raycast menu figure taken from [33]

with when you need to do a lot of tasks, a 3D interface greatly increases the amount of fun and presence in menu navigation [26]. Some research has been done into which type of 2D menu, out of many, works best in VR [33]. With panel menus scoring the highest. Fixing the menu in place, instead of having it tracked to an object such as your hand, is found to be more accurate, faster and preferred by the users [43]. Although extensive research has been done on menu design in 2D, research has yet to show that 2D design paradigms are also the most effective in a 3D VR environment where immersion and presence play greater parts. Overall there is a limited amount of research done into VR menu design.

2.4.1 No Controller

Some research has looked at ways of menu interaction outside of a 6-DoF controller or hand tracking. Sublime makes use of EEG data to read brain activity and select menu items [4]. Although it was slower than conventional methods, the user experience was better. Another study tried to supplement controller based UI with gaze interactions [34]. Gesture interaction is an interesting possibility as well, with many studies already taking a look at how to do this. The main drawback however is that gesture interactions are not as precise as a controller and gesture based systems have a much higher learning curve [48]. Most of these methods thus come with drawbacks such as interaction speed, more complex systems or require external devices which is not ideal if you want your application to be used as easily as possible.

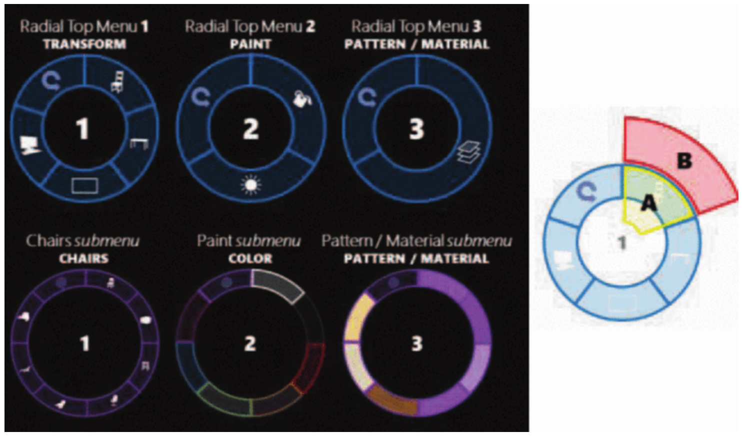

Figure 4: Multi-level radial menu, figure taken from [17]

2.4.2 Text Input

Another important part of menu interaction is text input. An empirical evaluation of four controller-based text-input techniques have found that the drum-like keyboard and the raycasting technique achieved higher usability and experiential feedback compared to head-directed input and a split keyboard [2]. Another interesting study used a physical smartphone that was tracked with the headsets cameras to be able to input text [5].

2.5 Literature Research Conclusion

When we look at existing literature we can clearly see that immersion affects learnability in educational VR applications. UI is part of the user experience and thus might also have an effect on immersion. However this was never researched directly. Using the current literature on how to achieve immersion in VR applications I will design a new method in the next section.

3 Method

As explained in section 2.3 it is important that my method includes some form of haptic feedback as that greatly increase the level of immersion. This has to be done without custom controllers. The officers of the Dutch police force make use of smartphones and smartwatches to interact with their systems during field operations. Based on their interaction design with those smartwatches I came up with a wrist tracked visual VR smart watch that can be tapped with the controllers. Tapping the smart watch will also tap the wrist of the user, thus creating haptic feedback of actually tapping something. A lot of people also have experience with using smart watches, which makes the design even more logical. I will also design a UI that visually simulates physical forces, as that also increase the presence experienced [8]. This section will explain the design of the smart watch and its variants. Then I will introduce my research questions and propose an experiment to answer those.

3.1 UIWatch Design

For the smart watch model itself, I will use the already existing 3D model from the PA. This is a simple black 3D smart watch model, with simplistic shading. It will be tracked to the hand that the user normally wears their watch on. With the face of the watch being on the top side of the wrist. See Figure 5 for a rendered image of the smart watch.

Figure 5: A render of the UIWatch model

3.2 UI Variants

There are three variants that each have an increasing amount of visual and haptic feedback. The first variant is chosen as it is the way most games do menus. the other variants are based on the literature found in the previous section. As shown in section 2.3, adding more feedback should increase the amount of immersion and presence experienced. Thus the following three UI variants is tested in my experiment:

3.2.1 Button Activated Raycast Fixed Menu (BRF)

This variant will be activated by clicking a button on the controller, which will bring up a menu that is placed in front of the user. The menu will be placed in the direction the user is looking, about 2 meters away. This menu does not follow any of the users movement. It can be interacted with by using the controllers as ray-guns, where aiming at a menu option and pulling the trigger will select it. This is the most used UI activation and interaction method in existing applications and thus will serve as a baseline.

3.2.2 UIWatch Activated Raycast Fixed Menu (URF)

This variant will be activated by tapping the wrist watch with the controller of the hand opposite to the UIWatch, which will bring up a menu that is placed in front of the user. The menu will be placed in the direction the user is looking, about 2 meters away. This menu does not follow any of the users movement. As with the aforementioned variant, it can be interacted with by using the controllers as ray-guns.

3.2.3 UIWatch Activated Touch Wrist Menu (UTW)

This variant will be activated by tapping the wrist watch with the controller of the hand opposite to the UIWatch. Doing so will bring up a menu that is tracked to the wrist of the user above the UIWatch. The menu can then be interacted with by touching the elements with the controller of the hand opposite to the UIWatch.

3.3 Evaluation Methods

To be able to reason about immersion and menu interaction we need two things: quantify the usability of the UI and measure the amount of immersion. For both I will be using the most relevant method that I also commonly found in other literature.

For the usability there are many options but the most commonly used method is the System Usability Scale (SUS) [12]. This questionnaire fits my research, as it is a quick and reliable way to determine system usability. There is no need for more complicated and accurate methods, as the focus of this research lies with the level of immersion created by the UIWatch. Yet it is still important to measure this because an unusable UI will only frustrate the user, no matter how immersive. Research has shown that a SUS score of 68 or higher means that the usability is above average. Thus the variants should score at least 68, otherwise they will be considered unusable.

Choosing the right way to evaluate the level of immersion is a little more difficult as a multitude of methods exist. I am going to discuss the most relevant methods I saw in the literature.

1. CAMIL [31]: The CAMIL framework is relatively new and tries to quantify the learnability of specifically HMD VR systems. The method is based on the assumption that media interacts with method. This framework includes a measure on immersion. However, it is made for complete applications while I am trying to test out a specific part of UI design. It also places more focus on learnability, while I think immersion and UI should be researched first.

2. GEQ [28]: The Game Experience Questionnaire is used in a number of papers that I found. It is a three module questionnaire that tries to asses a multitude of things. The most important module is the first module, the core questionnaire as it tests immersion. This questionnaire also tests challenge and competence which are useful for when you are testing an UI. It is possible to leave out the other modules and irrelevant questions as the questionnaire is designed to be modular in use. It is however a problematic questionnaire, as the original authors never published it and it never received any peer-review. A validation study also found problems with the consistency and deemed the whole method unstable [30]. Thus this method will not create scientifically sound results and will not be used in my experiment.

3. GIQ [15]: The Game Immersion Questionnaire is developed to be able to measure immersion in educational games with 24 questions in three modules based on the immersion theory of Brown and Cairns. The questionnaire was validated in the same paper it was published in and seems a reliable way of measuring immersion. However, I could not find out how to score the results as an explanation is not included in the paper. I also find the questions to not be completely suitable for what I want to measure.

4. PXI [24]: The Player Experience Inventory was designed to measure player experience in a variety of game genres and gamified applications. It is built on the Means-End theory and the related MDA framework. Immersion is a small but important part of this questionnaire for me. It also tests mastery, audiovisual appeal, ease of control and challenge which are nice supplements to the SUS questionnaire. The rest is irrelevant but needs to be included, otherwise the questionnaire is not validated. The PXI can also be found in a short version and is available in Dutch as well, which is useful since the application is tested at the Dutch Police Academy.

5. ARI [21]: The ARI questionnaire was specifically designed to measure immersion in AR location-based games. It however is not applicable for VR based UI testing, as the questions are geared towards activity based interactions with frequents distractions from real world stimuli. Something that will not happen in HMD based VR as you are completely shut off from the world around you.

Weighing the pros and cons of each method, I am going with the PXI. As that seems to be the most suitable method that is also validated. Both the SUS and PXI questionnaires will be made into a single Google Forms document, to be able to easily let the participants enter their answers.

3.4 Research Question

In this subsection I will introduce the research questions that I want to answer. These questions are a logical continuation from the literature research, as the answer to the question will contribute to a better understanding of UI and immersion in VR. The question is subdivided in three sub questions to be better able to answer the more complex main question.

1. Is UIWatch equally usable and more immersive than BRF?
    - (a) Are URF and UTW equally usable to BRF according to the SUS questionnaire?
    - (b) Is URF more immersive than BRF according to the PXI questionnaire?
    - (c) Is UTW more immersive than BRF according to the PXI questionnaire?
    - (d) Is UTW more immersive than URF according to the PXI questionnaire?

Hypotheses:

UIWatch is equally usable and more immersive than BRF.

- *H*1 Other research saw no decrease in competence and usability between touch and raycast. URF and UTW will not show a decrease in competence and usability comparaed to BRF.
- *H*2 Haptics is an important part of immersion, so the more involved method URF will be more immersive than BRF. This is because it has the haptic feedback added of touching your wrist to activate the menu.
- *H*3 As with URF, UTW has UIWatch activation. Thus it also creates more haptic feedback. Aside from that it also includes the visually simulated feedback of the Touch Wrist Menu. UTW will be more immersive than BRF.
- *H*4 UTW will be more immersive than URF as the UI interaction and placement seem to be more advantageous in creating visually simulated haptic feedback.

The questions and hypotheses contain the aspect of usability, which was not touched upon in the literature research as it is not the main focus of this study. Usability is however important to measure since a non-usable menu would only frustrate the user. It is also measured by most of the VR menu research mentioned previously. $H1$ is a logical continuation from what was found in the research from the previous section as they also did not see any decrease in competence and usability.

3.5 Experiment Design

In this subsection I will explain my experiment setup and how I am going to quantify the results and reason about them.

3.5.1 Materials

For the actual testing application I will use Unity [41]. For the headset, a Meta Quest 2 will be used alongside a Oculus Rift S. These headsets both run on the Oculus platform and have a similar controller design. Although the Meta Quest 2 has a higher resolution than the Oculus Rift S, both the text and the visual elements of the testing application are rather large and simple which means that the resolution difference does not translate into more detail. Thus I do not expect the headsets to produce different results as they show the exact same content. I will build the earlier mentioned UI variants into the test application to be able to run it on a VR headset. A floating panel about 2 meters away from the user will show the test instructions. The rest of the environment will stay empty as to not influence the results. See figure 6 for a screenshot of the environment and the instruction panel.

Figure 6: A screenshot of the test application showing the environment and floating instruction panel.

3.5.2 Menu Design

The design for the menu is very simple. There is one main menu (Figure 7) with 4 buttons that go to submenus. The four submenus are: Settings (Figure 8), Level Select (Figure 9), Language (Figure 10) and Exit Level (Figure 11). The menus are the same for both the Raycast Fixed Menu and the Touch Wrist Menu, however the Touch Wrist Menu has protruding buttons that can be pressed by the user with their hands (see Figure 12).

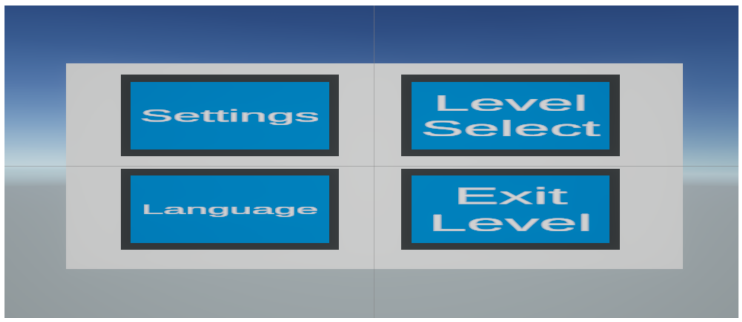

Figure 7: A screenshot of the Raycast Fixed main menu.

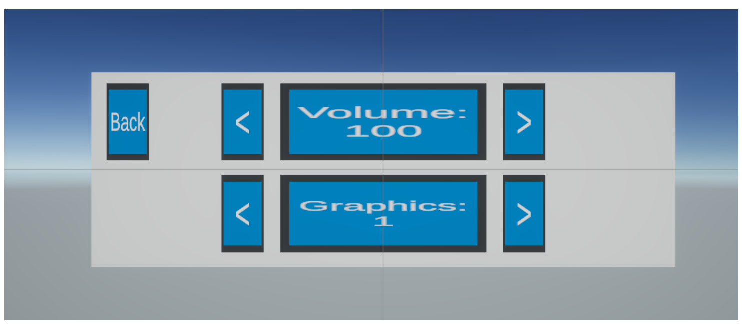

Figure 8: A screenshot of the Raycast Fixed Settings submenu.

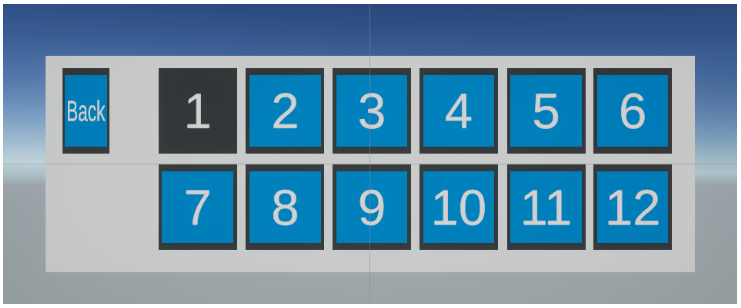

Figure 9: A screenshot of the Raycast Fixed Level Select submenu.

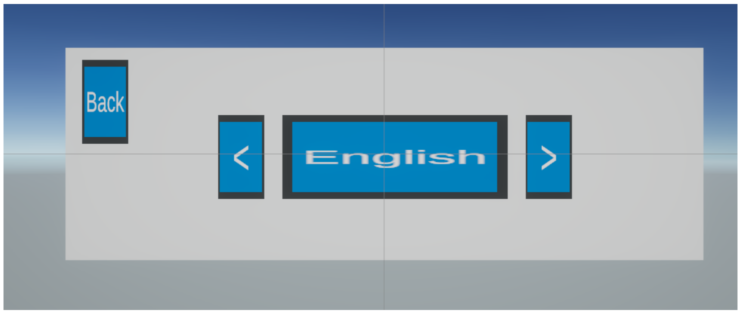

Figure 10: A screenshot of the Raycast Fixed Language submenu.

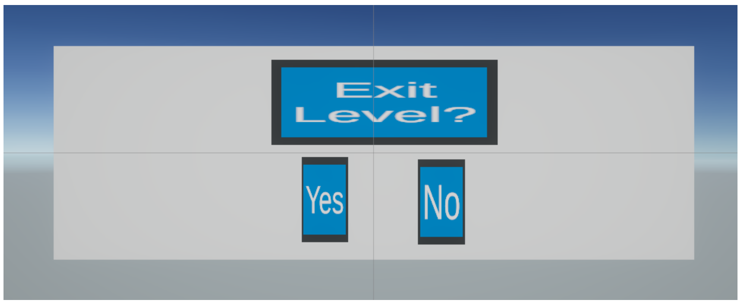

Figure 11: A screenshot of the Raycast Fixed Exit Level submenu.

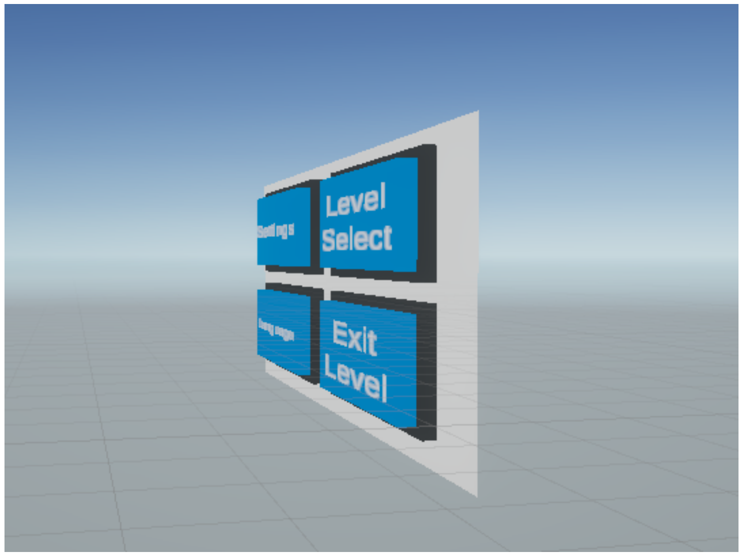

Figure 12: A screenshot of the Touch Wrist main menu with protruding buttons.

3.5.3 Test Procedure

The experiment will be within-subject. Where each participant will test out all three UI variants in a set order. In total at least 20 participants are needed across both the pilot and full study to verify the menu variants. This will ensure that the results are reliable and can be generalized to a larger population. The test participants will first be asked to digitally sign the consent form made by Utrecht University [1]. Before the participants start testing, they will get the chance to get familiar with the menu layout. This will eliminate the learning effect since the participants already know where everything is. Then they will do 5 UI tasks with each of the three variants (in this order: BRF, UTW, URF) and after each variant they will fill in the SUS and PXI questionnaires outside of the VR environment. The questionnaires will be filled in outside of the VR environment to make sure there is no carryover effect in terms of immersion. These tasks will also be timed from when the task is shown until it is completed, these task time results will then be used to compare the speed of each UI. UI speed is important to measure since a menu should not take unnecessarily long to navigate. Task time will in this case thus be used to measure competence. Since the tasks are all different in terms of interactions and where the buttons are placed, it is important to look at each task individually as well as the average task time. If you only look at the average you will miss it if one of the tasks is more difficult in one of the variants. The UI tasks are the following:

1. **Volume**: In the Settings submenu, set the volume to 75.
2. **Level**: In the Level Select submenu, select level 7.

3. **Language**: In the Language submenu, set the language to Italiano.
4. **Graphics**: In the Settings submenu, set the graphics to 4.
5. **Exit**: In the Exit Level submenu, exit the level by pressing yes.

To summarize the test procedure will look like this:

1. Sign the consent form on a laptop
2. Use and learn the menu layout until the participant feels comfortable and competent to carry out the tasks
3. Perform the 5 UI tasks with the BRF variant inside the VR environment.
4. Fill out the SUS and PXI questionnaire about the BRF variant on a laptop outside the VR environment.
5. Perform the 5 UI tasks with the UTW variant inside the VR environment.
6. Fill out the SUS and PXI questionnaire about the UTW variant on a laptop outside the VR environment.
7. Perform the 5 UI tasks with the URF variant inside the VR environment.
8. Fill out the SUS and PXI questionnaire about the URF variant on a laptop outside the VR environment.

3.6 Result Processing

For the task time results, one group tested all three variants. There is only one independent variable and the data is measured at the continuous level. A possible test would be the one way ANOVA with repeated measures to analyze the task time data [3]. The assumptions of this ANOVA test are as follows:

1. Your dependent variable should be measured at the continuous level: the task time data is continuous.
2. Your independent variable should consist of at least two categorical, related groups or matched pairs: the three variants are tested by the same group of people.
3. There should be no significant outliers in the related groups: this is tested for when the data is ran through SPSS for analysis.
4. The distribution of the dependent variable in the two or more related groups should be approximately normally distributed: this will also be tested for when the data is ran through SPSS.
5. The variances of the differences between all combinations of related groups must be equal: a sphericity check is built into the SPSS ANOVA tool. If this assumption fails a greenhouse-geisser correction will be applied as explained later on in this paragraph.

This test method is used to see if there is a difference between the task times of the three variants. ANOVA however relies on the assumption that the data is "spherical". The variance of the differences between all combinations must be equal. If this is not the case a greenhouse-geisser correction needs to be applied. This correction applies some extra math to the calculation of the p-value, which compensates for the violation of sphericity.

The PXI and SUS rating scale data are ordinal, which means the distance between the scores is not the same. For example the difference between a 5 and a 6 for a question on the PXI questionnaire is not the same as the distance between a 6 and a 7. This ordinality also means a non-parametric analysis method is needed. One group, randomly sampled from the population, is measured on three different occasions for both questionnaires. A possible test would be the Friedman test method [19]. The assumptions of the Friedman test are as follows:

1. One group is measured on three or more different occasions: the three variants are tested by the same group of people.
2. Group is a random sample from the population: there were no selection criteria besides the participants having to be 18 or older.
3. Your dependent variable should be measured at the ordinal or continuous level: the PXI and SUS data is ordinal.
4. Samples do NOT need to be normally distributed: this assumption always succeeds.

The data matches all assumptions of the Friedman test method, thus the data will be analyzed using this method. This test method is similair to the one-way ANOVA mentioned in the previous paragraph. As with the ANOVA, it is also used to test for differences between the variants. If a difference is found in a Friedman test, post hoc tests on all three pairings should be done using the Wilcoxon signed-rank test analysis [45]. This Wilcoxon test compares each variant to the other two variants and checks if the difference between them is significant. All these tests will result in a p-value, which shows how certain you can be that the results differ from each other. This p-value is compared to an alpha value that you decide on beforehand. The alpha value of this research is set to 0.05 since that is the most common value. So if the p-value is lower than 0.05 you can see the results as significantly different. The Wilcoxon tests are making multiple comparisons, which increases the chance of declaring a result significantly different when it is not. Thus we need to adjust the alpha value to avoid making this error. The Bonferroni adjustment can be used for this. The Bonferroni adjustment is calculated by taking the original alpha value 0.05 and dividing it by the amount of variants that were tested, which is 3. This means that for the Wilcoxon tests an alpha of 0.017 should be used.

4 Results

This section will contain the results of both the pilot study and the experiment.

4.1 Pilot Study

The first round of testing was done with the VR team of the PA Leusden. These tests were done in a relatively quiet room with the Quest 2 headset. In total 5 participants were part of this round with an average age of $27.4(SD : 4.3)$ and a median VR experience of over 100 hours total played.

	BRF		**URF**		**UTW**	
Metric	Mean	SD	Mean	SD	Mean	SD
SUS	79.5	8.91	63.5	16.36	76.5	12.45
Immersion (PXI)	11.2	3.27	12.6	4.04	14.2	4.44
Audiovisual Appeal (PXI)	8.2	3.11	10.6	5.03	11.8	6.30
Ease of Control (PXI)	17	2.45	15.6	2.61	17	4.06

Table 1: Mean and standard deviation of all pilot study questionnaire feature values of the evaluated UI VR variants.

Table 1 shows the results of the SUS and PXI questionnaires. Both BRF and UTW scored above 68, which is considered the cutoff between usable UI and UI that needs improvement. URF scored below 68, which means it is below average. For the PXI results, each score is out of a maximum of 21. All variants scored around 12.5 on immersion, with UTW scoring the highest (14.2). Audiovisual Appeal is scored a little lower than immersion but UTW again scored the highest. Ease of control was scored relatively high with BRF and UTW scoring equal.

	BRF		**URF**		**UTW**	
Metric	Mean	SD	Mean	SD	Mean	SD
Volume	11.8	3.92	8.1	1.69	10.9	2.67
Level	5.7	2.00	4.7	0.52	7.2	1.84
Language	5.0	1.31	4.7	1.41	6.0	1.43
Graphics	6.1	1.24	6.0	1.53	7.2	1.19
Exit	5.3	1.13	3.3	0.70	5.6	2.44
Average	6.8	1.14	5.4	0.59	7.4	0.76

Table 2: Mean and standard deviation of all pilot study timed tasks in seconds.

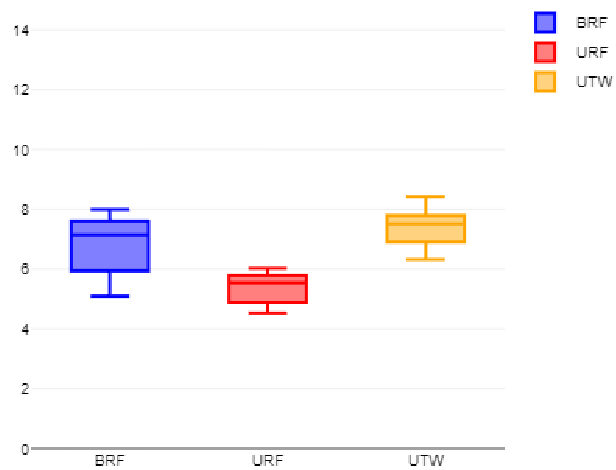

Figure 13: A boxplot of the average task time results for the pilot study.

Table 2 shows the results of the task times. Task times differ both between variants and between tasks. URF was the fastest with 5.4, followed by BRF with 6.8 and UTW with 7.4. Note that the standard deviations are high due to the limited amount of participants. These results can also be seen as a boxplot in Figure 13.

	χ^2	**df**	**P-Value**
SUS	1.600	2	0.449
Immersion	5.261	2	0.05
Audiovisual Appeal	4.667	2	0.097
Ease of Control	1.333	2	0.549

Table 3: Results of pilot study questionnaire Friedman test analysis with alpha = 0.05.

Table 3 shows there was no statistically significant difference in how people score the different UI variants.

	F	df	P-Value
Volume	2.905	2	0.113
Level	4.829	2	0.042
Language	1.306	2	0.323
Graphics	1.013	2	0.405
Exit	4.121	2	0.059
Average	7.185	2	0.016

Table 4: Results of pilot study task ANOVA analysis with alpha = 0.05, the Bonferroni adjustment has already been applied on the numbers in the table. The sphericity assumption was violated nowhere and thus not reported.

First Type	Second Type	P-Value
BRF	UTW	0.445
BRF	URF	0.699
UTW	URF	0.133

Table 5: Pairwise comparison of the pilot study level task time results.

First Type	Second Type	P-Value
BRF	UTW	1.000
BRF	URF	0.041
UTW	URF	0.154

Table 6: Pairwise comparison of the pilot study exit task time results.

First Type	Second Type	P-Value
BRF	UTW	1.000
BRF	URF	0.222
UTW	URF	0.003

Table 7: Pairwise comparison of the pilot study average task time results.

A repeated measures ANOVA determined that average task time differed statistically significantly between the different types ($F(2,8) = 7.185, P < 0.05$). Post hoc analysis with a Bonferroni adjustment revealed that URF was statistically significantly faster than UTW. See table 4 and table 7. There were significant differences in the level (Figure 5, exit (Figure 6 and average (Figure 7 task times.

4.2 Discussion Pilot Study

Since the number of participants is very limited, it is not possible to draw any conclusions from this data. However, participants did show a decrease in task time during the test session. The participants get time to get familiar with the menu layout before the tests start, so there should not be a significant learning effect there. However, it could be due to the participants learning the task order during the test. To combat this possible learning effect, the task order of UTW and URF have been shuffled for the next test round to make them less predictable.

It is also interesting to note that although the difference for immersion is still insignificant, the P-Value is close to 0.05. This suggests that the design might be more immersive and that the it is worth exploring further with the full experiment.

4.3 Full Study

The full study contained 21 participants with an mean age of 33 (SD: 13.9), which existed out of 18 male participants and 3 female participants. 2 of these participants were also part of the Pilot study, the others had never seen the test application before. There was no significant difference in the results of the participants that used the application before. The reason for that might be the full month between the pilot study and full study. However, they might have had faster task times if the studies were closer to each other. 16 participants reported having experienced more than 10 hours of VR. Both the Meta Quest 2, with 11 participants, and the Oculus Rift 2, with 10 participants, were used.

	BRF		**URF**		**UTW**	
Metric	Mean	SD	Mean	SD	Mean	SD
SUS	80.1	11.17	69.04	19.31	80.58	16
Immersion	14	2.84	14.85	4.4	16.77	4.02
Audiovisual Appeal	10.77	4.54	12.92	4.69	14.92	5.22
Ease of Control	17.73	2.11	15.88	3.39	17.77	2.86

Table 8: Mean and standard deviation of all full study questionnaire feature values of the evaluated UI VR variants.

Table 8 shows the results of the SUS and PXI questionnaires for the full study. All three UI variants scored above 68, with UTW scoring the highest. All variants scored 14 or higher on immersion which is above average, with UTW scoring the highest (16.77). Audiovisual Appeal is scored lower than immersion but UTW again scored the highest. Ease of control was scored relatively high with BRF and UTW scoring almost equal.

	BRF		URF		UTW	
Metric	Mean	SD	Mean	SD	Mean	SD
Volume	10.16	3.96	8.76	3.73	9.16	3.14
Level	8.18	5.49	5.22	2.82	7.69	4.53
Language	6.53	2.61	5.59	2.22	6.51	1.99
Graphics	7.86	3.16	5.55	3.01	8.35	2.61
Exit	5.78	3.96	7.14	3.77	5.48	2.49
Average	6.91	2.55	6.4	1.69	8.99	2.57

Table 9: Mean and standard deviation of all full study timed tasks in seconds.

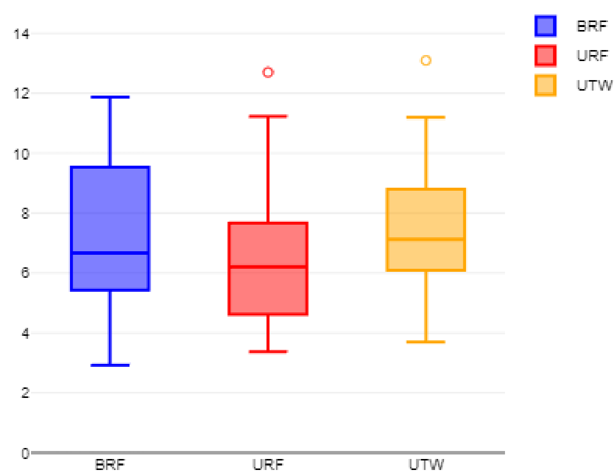

Figure 14: A boxplot of the average task time results for the full study.

Table 9 shows the results of the task times for the full study. As with the pilot study, task times differ both between variants and between tasks. URF again was the fastest with 6.4, followed by BRF with 6.91 and UTW with 8.99. These results can also be seen as a boxplot in Figure 14.

	χ^2	**df**	**P-Value**
SUS	3.325	2	0.190
Immersion	13.486	2	0.001
Audiovisual Appeal	23.077	2	< 0.001
Ease of Control	4.630	2	0.099

Table 10: Results of full study questionnaire Friedman test analysis with alpha = 0.05.

First Type	**Second Type**	**P-Value**
BRF	UTW	< 0.001
BRF	URF	0.341
UTW	URF	0.005

Table 11: Pairwise Wilcoxon signed rank test comparison of the full study immersion PXI results.

First Type	**Second Type**	**P-Value**
BRF	UTW	< 0.001
BRF	URF	0.009
UTW	URF	0.008

Table 12: Pairwise Wilcoxon signed rank test comparison of the full study audiovisual appeal PXI results.

Table 10 shows that there was a statistically significant difference between how the UI variants were rated on Immersion and Audiovisual Appeal ($P = 0.001$, $P < 0.001$). Post hoc analysis with Wilcoxon signed-rank tests on Immersion, see table 11, showed that UTW was scored higher than both BRF and URF. Post hoc analysis with Wilxocon signed-rank tests on Audiovisual Appeal, see table 12, showed that UTW was scored the highest following URF.

	P-Value	**F**
Volume	0.210	1.623
Level	0.020	4.962
Language	0.128	2.166
Graphics	< 0.001	8.748
Exit	0.160	1.922
Average	0.051	3.199

Table 13: Results of full study task ANOVA analysis with alpha = 0.05. The sphericity assumption was only violated with the level results (sig. = 0.030) for which the greenhouse-geisser correction was used ($df = 1.529$ $df(error) = 30.582$).

First Type	Second Type	P-Value
BRF	UTW	1.000
BRF	URF	0.002
UTW	URF	0.062

Table 14: Pairwise comparison of the full study level task time results.

First Type	Second Type	P-Value
BRF	UTW	1.000
BRF	URF	0.023
UTW	URF	0.002

Table 15: Pairwise comparison of the full study graphics task time results.

A repeated measures ANOVA determined that average task time did not differ statistically significantly between the different types ($F(2, 40) = 3.199, P = 0.051$). Although the p-value is really close to 0.05.

5 Discussion

This section will contain the discussion of the results of the full study and the limitations.

5.1 Participant demographics

In total 24 participants participated in this study. This is more than the needed number of 20 that was determined beforehand. The participants were spread in terms of VR experience with about a third having very little experience (¡10 hours), a third having played a decent amount (10-100 hours) and the last third having played a lot (100+). These three groups did not perform, on average, significantly different on the timed tasks and rated the menu variants equally. Thus all participants were regarded as one group in the analysis.

5.2 Task Time Results

The task time results from the full study, which can be seen in table 9, did not show a significant difference between each variant. Although some of the individual task times were significantly different, it is not enough to say that one variant is quicker or slower than any other. This is similar to the results from earlier literature that compared different UI designs, such as the research done by Monteiro et al. [33] and Davis et al. [17] that were mentioned in the literature research section. Their research also did not see any difference in task times between their menu variants. The results are also in line with what was expected based on the experiment design and literature research. However, since the P-value was very close to the alpha, it is still interesting to take a closer look at the results and what might have caused this. Interestingly URF, although being reported by the users as the least easy to use, seemed to be the fastest. This might be due to the fact that when you activate the menu with your hand, you are already pointing forward with the controller. Thus the raycast will already be pointing at where the menu appears, decreasing aiming time in comparison with BRF. UTW is about 2 seconds slower than the other two, which might be because you have to move your arms around to select menu items compared to just pointing at them. This is however speculation and a future study might be needed to look into hand movements and task time in more detail.

5.3 Questionnaire Results

5.3.1 SUS

Each variant was scored above 68 on the SUS questionnaire, which means that each variant was scored above average on usability. The SUS scores of the full study were a bit higher than the SUS scores from the pilot study. This might be due to the fact that the pilot study participants consisted out of mostly people that work with VR, while the full study was a more random sample of the population in terms of VR experience. The people that work with VR have seen more VR menus, with a high chance that one of those menus was designed better and thus they will rate anything worse than their best viewed VR menu as worse. Whereas the people that have less VR experience do not have a frame of reference to other VR menus and thus might rate everything a bit higher on the SUS questionnaire. In hindsight, it would have been a good idea to interview the people that had a lot of VR experience during the pilot study about possible improvements to the menu. These improvements could then have been implemented in the full study to possibly increase the SUS scores.

Comparing these results to the SUS results from Monteiro et al. [33]; BRF, URF and UTW scored 10 to 20 points lower than their menu and interaction variants. This might be due to the fact that in their study the menu controlled certain things in the environment such as turning on the lights, while in this study the menu did nothing. Maybe the link between menu action and environmental consequence made the participants rate the menu higher on usability. This idea is supported by the findings of Sun et al. [40], which found that more direct manipulation leads to better performance. Better performance might then in turn lead to a higher rating for usability. The menu of Monteiro was also more aesthetic, with the inclusion of things like pictograms for buttons. A study done by Hartman et al. [25] showed that aesthetics have an influence on UI usability. Those factors might have had an influence on the usability ratings for their study.

There was no significant difference between the three variants. URF did score about 10 points lower than BRF and UTW. URF was always tested last in the full study. Participants might have seen the URF variant as a mix between BRF and UTW since it is comprised of elements from BRF (the raycast menu) and

UTW (the smart watch activation method). Participants could have thought that it was not an improvement over those two as it added nothing new. Showing URF in isolation might get it scored higher but then it is up to discussion which of those scores can be trusted more.

5.3.2 PXI

Each of the variants scored high on Ease of Control, with no significant difference between them. This suggests that the variants are all easy to use. As with the SUS and task time results, the same literature [33] [17] supports this observation. These results also fit in with the results from the task times and SUS questionnaire, as all three show that all variants are both usable and make the user feel competent.

All three variants scored at least a 14 out of 21 on Immersion, using the scale from the PXI this means that the users reported that they felt somewhat immersed. Post hoc tests on the PXI immersion results showed that UTW, with a high mean score of 16.77, was significantly more immersive than BRF and URF. The score of 16.77 suggests that users felt greatly immersed during their use of the UTW variant. This result coincides with what was expected based on the section "How to achieve immersion" from the literature research. It is also supported by what was found by Biocca et al. during their study into cross-modal visual-to-haptic transfers [8]. The immersion results from their more visually simulated haptic variant are in line with what was found in this study. They found that if you visually simulate haptics, users will report higher levels of immersion. Which is exactly the case in this study as well.

URF however, was not more immersive than BRF while the literature research did suggest there should be a difference. As mentioned earlier for the SUS scores, this might be caused by showing URF last during the experiment. URF is also very similar to BRF with the only difference being their menu activation method. Most of the time during the experiment was spent in the menu, as opposed to activating it, so they might have felt very similar to the user as well. This could explain the similar immersion scores.

Audiovisual appeal was scored relatively low compared to Immersion and Ease of Control. This might be because the UI was designed for research on immersion based on earlier research and not for aesthetics. UTW was still scored the highest (14.92) with a significant difference to BRF and URF. This might be because actually having to click the virtual button with your hand for UTW was novel and fun for most participants. However, to be sure of the particular reason for this difference a follow up study with qualitative questions about the aduiovisual appeal is needed.

5.4 Discussion Summary

To summarize, participants favored UIWatch Activated Touch Wrist Menu (UTW) over both Button Activated Raycast Fixed Menu (BRF) and UIWatch Activated Raycast Fixed Menu (URF). UTW was scored the most immersive and did not differ in terms of usability and competence, measured in task time, from BRF and URF. Although UTW was the clear favorite, the results show all three variants are usable as they all scored above 68 on the SUS questionnaire and were rated high on Ease of Use.

5.5 Limitations

This subsection will discuss the limitations of this research:

1. The tests were only done on the Oculus platform with their touch controller design. Other platforms use other types of controllers which allow more interactions such as letting the controller go fully with the Valve Index controllers. These interactions were not considered in this research.

2. The UI variants were tested in a stripped down environment and users were forced to perform tasks. This is not how a menu would be normally used in an application. It would be interesting to see how the menus would perform in more complete VR applications as the results might differ when using the variants for a longer amount of time. After a longer period of usage people might not be interested in the more involved method of UTW as it requires more effort. This can affect the immersion since users might get annoyed by the menu instead of immersed by how involved it is. A longer study in the future might explore this in more depth.

3. Extensive literature research guided the designs and interaction mechanics of the variants, however the aesthetics of the UI were not the main focus. With more focus on aesthetics, by an experienced designer, the variants might score even higher on immersion. This idea is supported by Bowman et al. [9], who state that visual immersion is affected by how close the system's visual output is to real-world visual stimuli. A more realistically designed menu might thus invoke higher levels of immersion. This however needs to be confirmed by a future study as there is no existing literature that combines menu design, haptics, aesthetics and immersion in VR.

6 Conclusion

Going back to the hypotheses we can conclude that $H1$, there will be no decrease in competence and usability between touch and raycast, can be accepted. $H2$, URF will be more immersive than BRF, has to be rejected. While URF was not more immersive than BRF, UTW was. This means that we can accept $H3$, which stated that UTW will be more immersive than BRF. $H4$, UTW will be more immersive than URF, can also be accepted. In conclusion, adding visually simulated haptics to menu navigation increases immersion without affecting usability. Adding haptics to the activation of a menu does not seem to increase immersion, however it also does not decrease immersion either. As other research has shown, more immersion leads to increased learnability in educational

VR applications. As this research shows it is advised to use simulated haptics, such as the UIWatch, over no haptics at all when making an educational VR application.

References

[1] Sept. 2022. URL: https://www.uu.nl/en/research/institute-of-information-and-computing-sciences/ethics-and-privacy.

[2] Jiban Adhikary and Keith Vertanen. "Text entry in virtual environments using speech and a midair keyboard". In: *IEEE Transactions on Visualization and Computer Graphics* 27.5 (2021), pp. 2648–2658.

[3] *Anova test in SPSS statistics*. URL: https://statistics.laerd.com/spss-tutorials/one-way-anova-repeated-measures-using-spss-statistics.php.

[4] Alexandre Armengol-Urpi and Sanjay E Sarma. "Sublime: a hands-free virtual reality menu navigation system using a high-frequency SSVEP-based brain-computer interface". In: *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*. 2018, pp. 1–8.

[5] Huidong Bai et al. "Bringing full-featured mobile phone interaction into virtual reality". In: *Computers & Graphics* 97 (2021), pp. 42–53.

[6] Mehmet Ilker Berkman and Ecehan Akan. *Presence and Immersion in Virtual Reality*. 2019.

[7] Svetlana Bialkova and Marnix S Van Gisbergen. "When sound modulates vision: VR applications for art and entertainment". In: *2017 IEEE 3rd Workshop on Everyday Virtual Reality (WEVR)*. IEEE. 2017, pp. 1–6.

[8] Frank Biocca et al. "Visual cues and virtual touch: Role of visual stimuli and intersensory integration in cross-modal haptic illusions and the sense of presence". In: *Proceedings of presence* (2002), pp. 410–428.

[9] Doug A Bowman and Ryan P McMahan. "Virtual reality: how much immersion is enough?" In: *Computer* 40.7 (2007), pp. 36–43.

[10] William Bricken. *Learning in Virtual Reality*. 1990.

[11] William Bricken. "Virtual reality: Directions of growth notes from the SIGGRAPH'90 panel". In: *Virtual Reality: Directions of Growth* 16 (1990).

[12] John Brooke et al. "SUS-A quick and dirty usability scale". In: *Usability evaluation in industry* 189.194 (1996), pp. 4–7.

[13] David Checa and Andres Bustillo. "A review of immersive virtual reality serious games to enhance learning and training". In: *Multimedia Tools and Applications* 79 (2020), pp. 5501–5527.

[14] Shengnan Chen et al. "A case study of user immersion-based systematic design for serious heritage games". In: *Multimedia Tools and Applications* 62 (2013), pp. 633–658.

[15] M-T Cheng, H-C She, and Leonard A Annetta. "Game immersion experience: its hierarchical structure and impact on game-based science learning". In: *Journal of computer assisted learning* 31.3 (2015), pp. 232–253.

[16] Robert Coulter et al. “The effect of degree of immersion upon learning performance in virtual reality simulations for medical education”. In: *In-Medicine Meets Virtual Reality* 15 (2007), p. 155.

[17] Matthew M Davis et al. “Depth-based 3D gesture multi-level radial menu for virtual object manipulation”. In: *2016 IEEE Virtual Reality (VR)*. IEEE. 2016, pp. 169–170.

[18] Laura Freina and Michela Ott. “A literature review on immersive virtual reality in education: state of the art and perspectives”. In: *The international scientific conference elearning and software for education*. Vol. 1. 133. 2015, pp. 10–1007.

[19] *Friedman test in SPSS statistics*. URL: `https://statistics.laerd.com/spss-tutorials/friedman-test-using-spss-statistics.php`.

[20] Eisuke Fujinawa et al. “Computational design of hand-held VR controllers using haptic shape illusion”. In: *Proceedings of the 23rd acm symposium on virtual reality software and technology*. 2017, pp. 1–10.

[21] Yiannis Georgiou and Eleni A Kyza. “The development and validation of the ARI questionnaire: An instrument for measuring immersion in location-based augmented reality settings”. In: *International Journal of Human-Computer Studies* 98 (2017), pp. 24–37.

[22] Janet K Gibbs, Marco Gillies, and Xueni Pan. “A comparison of the effects of haptic and visual feedback on presence in virtual reality”. In: *International Journal of Human-Computer Studies* 157 (2022), p. 102717.

[23] Alessandra Gorini et al. “The role of immersion and narrative in mediated presence: the virtual hospital experience”. In: *Cyberpsychology, behavior, and social networking* 14.3 (2011), pp. 99–105.

[24] Aqeel Haider, Kathrin Gerling, and Vero Vanden Abeele. “The Player Experience Inventory Bench: Providing Games User Researchers Actionable Insight into Player Experiences”. In: *Extended Abstracts of the 2020 Annual Symposium on Computer-Human Interaction in Play*. 2020, pp. 248–252.

[25] Jan Hartmann, Alistair Sutcliffe, and Antonella De Angeli. “Towards a theory of user judgment of aesthetics and user interface quality”. In: *ACM Transactions on Computer-Human Interaction (TOCHI)* 15.4 (2008), pp. 1–30.

[26] Daniel Hepperle et al. “2D, 3D or speech? A case study on which user interface is preferable for what kind of object interaction in immersive virtual reality”. In: *Computers & Graphics* 82 (2019), pp. 321–331.

[27] Hsiu-Mei Huang, Ulrich Rauch, and Shu-Sheng Liaw. “Investigating learners’ attitudes toward virtual reality learning environments: Based on a constructivist approach”. In: *Computers & Education* 55.3 (2010), pp. 1171–1182.

[28] Wijnand A IJsselsteijn, Yvonne AW De Kort, and Karolien Poels. “The game experience questionnaire”. In: *Not Published* (2013).

[29] Lasse Jensen and Flemming Konradsen. “A review of the use of virtual reality head-mounted displays in education and training”. In: *Education and Information Technologies* 23.4 (2018), pp. 1515–1529.

[30] Effie L-C Law, Florian Brühlmann, and Elisa D Mekler. “Systematic review and validation of the game experience questionnaire (geq)-implications for citation and reporting practice”. In: *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*. 2018, pp. 257–270.

[31] Guido Makransky and Gustav B Petersen. “The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality.” In: *Educational Psychology Review* (2021).

[32] Fabrizia Mantovani et al. “Virtual reality training for health-care professionals”. In: *CyberPsychology & Behavior* 6.4 (2003), pp. 389–395.

[33] Pedro Monteiro et al. “Comparison of radial and panel menus in virtual reality”. In: *IEEE Access* 7 (2019), pp. 116370–116379.

[34] Ken Pfeuffer et al. “Empirical evaluation of gaze-enhanced menus in virtual reality”. In: *26th ACM symposium on virtual reality software and technology*. 2020, pp. 1–11.

[35] Jaziar Radianti et al. “A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda”. In: *Computers & Education* 147 (2020), p. 103778.

[36] Michael A Rupp et al. “Investigating learning outcomes and subjective experiences in 360-degree videos”. In: *Computers & Education* 128 (2019), pp. 256–268.

[37] Yuki Sawada et al. “Effects of synchronised engine sound and vibration presentation on visually induced motion sickness”. In: *Scientific reports* 10.1 (2020), pp. 1–10.

[38] Mel Slater. “A note on presence terminology”. In: *Presence connect* 3.3 (2003), pp. 1–5.

[39] Mel Slater. “Immersion and the illusion of presence in virtual reality”. In: *British journal of psychology* 109.3 (2018), pp. 431–433.

[40] Chengyu Sun, Wei Hu, and Diqiong Xu. “Navigation modes, operation methods, observation scales and background options in UI design for high learning performance in VR-based architectural applications”. In: *Journal of Computational Design and Engineering* 6.2 (2019), pp. 189–196.

[41] Unity Technologies. URL: `https://unity.com/`.

[42] Hakan Tüzün and Fatih Özdinç. “The effects of 3D multi-user virtual environments on freshmen university students’ conceptual and spatial learning and presence in departmental orientation”. In: *Computers & Education* 94 (2016), pp. 228–240.

[43] Yanbin Wang, Yizhou Hu, and Yu Chen. "An experimental investigation of menu selection for immersive virtual environments: fixed versus handheld menus". In: *Virtual Reality* 25.2 (2021), pp. 409–419.

[44] Michael White et al. "The case for haptic props: Shape, weight and vibrotactile feedback". In: *Proceedings of the 12th ACM SIGGRAPH Conference on Motion, Interaction and Games*. 2019, pp. 1–10.

[45] *Wilcoxon test in SPSS statistics*. URL: https://statistics.laerd.com/spss-tutorials/wilcoxon-signed-rank-test-using-spss-statistics.php.

[46] Haojie Wu et al. "3d sound rendering in a virtual environment to evaluate pedestrian street crossing decisions at a roundabout". In: *2018 IEEE 4th VR Workshop on Sonic Interactions for Virtual Environments (SIVE)*. IEEE. 2018, pp. 1–6.

[47] Biao Xie et al. "A review on virtual reality skill training applications". In: *Frontiers in Virtual Reality* 2 (2021), p. 645153.

[48] LI Yang et al. "Gesture interaction in virtual reality". In: *Virtual Reality & Intelligent Hardware* 1.1 (2019), pp. 84–112.

[49] Christine Youngblut. "Educational uses of virtual reality technology." In: (1998).

[50] Andre Zenner and Antonio Krüger. "Shifty: A weight-shifting dynamic passive haptic proxy to enhance object perception in virtual reality". In: *IEEE transactions on visualization and computer graphics* 23.4 (2017), pp. 1285–1294.