USER EXPERIENCE GUIDELINES FOR DESIGN OF VIRTUAL REALITY GRAPHICAL USER INTERFACES controlled by head orientation input

Bachelor Degree Project in Cognitive Science
Three years Level 30 ECTS
Spring term 2016

Sofia Fröjdman

Supervisor: Tarja Susi
Examiner: Niklas Torstensson
Abstract

With the recent release of head-mounted displays for consumers, virtual reality experiences are more accessible than ever. However, there is still a shortage of research concerning how to design user interfaces in virtual reality for good experiences. This thesis focuses on what aspects should be considered when designing a graphical user interface in virtual reality - controlled by head orientation input - for a qualitative user experience. The research has included a heuristic evaluation, interviews, usability tests, and a survey. A virtual reality prototype of a video on demand service was investigated and served as the application for the research. Findings from the analysis of the data were application specific pragmatic and hedonic goals of the users, relevant to the subjective user experience, and current user experience problems with the prototype tested. In combination with previous recommendations, the result led to the development of seven guidelines. However, these guidelines are considered only to serve as a foundation for future research since they need to be validated. New head-mounted displays and virtual reality applications are released every day and with the increasing number of users, there will be a continuous need for more research.

Keywords: Virtual reality graphical user interface, Head orientation input, Head-mounted display, User experience, Video on demand service
Popular scientific abstract

As a result of the recent release of head-mounted displays and virtual reality sets, such as Oculus Rift, Samsung Gear VR, and Google Cardboard, virtual reality has turned into a big technology trend. Since the virtual reality experiences just recently have begun to become available for consumers, there is still a lack of research concerning the users experience of them. How should user interfaces be designed for virtual reality to be easy and fun to use and how should the user interface be controlled in the best possible way? This thesis focuses on the only available control method that is possible to use with all different head-mounted displays and virtual reality sets. The way the user with a head-mounted display turn his or her head in different directions, allow the user to see more of the virtual world, and this could also be used to project a gaze cursor in the centre of the user’s field of view. If the user then stays with the gaze cursor over a certain object in the user interface for a predetermined period of time, a selection can be made. This thesis focuses on answering the question of what aspects should be considered when designing a user interface in virtual reality, controlled in this way, for qualitative user experiences. A prototype application of a video on demand service in virtual reality was used for part of the research and tested by users. Additionally, users were interviewed and asked in a survey about their goals, knowledge, preferences, and expectations. The result of the research showed the goals of why users wanted to use a video on demand service in virtual reality with a head-mounted display and what they wanted to be able to do in such an application. It also showed current user experience problems with the application that was tested. Based on these results, together with previous recommendations found in literature, new guidelines were developed. However, these need to be tested and further confirmed before they can be used. With the release of new head-mounted displays and virtual reality applications every day, there will be a growing number of user, and this will lead to an increasing need for research about how to improve the experiences as well.
Acknowledgements

Writing this thesis has been a challenging but very exciting process. Moreover, it has been a tremendous learning experience, and I am very grateful for all the inspiring and supportive people around me who have helped me along the process. Foremost, I would like to express my sincere gratitude to my supervisor Tarja Susi for all the continuous support and guidance during these months. I would also like to thank Niklas Torstensson for his feedback and encouragement.

Moreover, I would like to thank people at Accedo for welcoming and supporting me as well as sharing their knowledge. I specifically want to thank my Accedo supervisor, José Somolinos, who has inspired me and introduced me to the world of virtual reality.

My sincere thanks also go to all the participants in the research who has taken their time to help me. Last but not least, I would like to thank my family and friends for their never-ending encouragement and support.
# Contents

List of abbreviations .................................................................................................................. 1

1 Introduction .................................................................................................................................. 1

2 Background .................................................................................................................................. 3
  2.1 Virtual reality and related concepts ....................................................................................... 3
  2.2 Virtual reality graphical user interface ................................................................................... 5
  2.3 Head-mounted display ............................................................................................................. 7
  2.4 Head orientation input and head gaze ..................................................................................... 9
  2.5 User experience ..................................................................................................................... 11
  2.6 Video on demand service ....................................................................................................... 12
  2.7 Research related design principles for virtual reality graphical user interfaces ................. 13
  2.8 Research aim and objective .................................................................................................. 16
    2.8.1 Expected contributions .................................................................................................... 17
    2.8.2 Limitations ..................................................................................................................... 17

3 Method ....................................................................................................................................... 18
  3.1 Research design ..................................................................................................................... 18
  3.2 Heuristic evaluation .............................................................................................................. 19
  3.3 Interview ............................................................................................................................... 21
  3.4 Usability test ........................................................................................................................ 22
  3.5 Survey ................................................................................................................................... 24

4 Procedure ................................................................................................................................... 25
  4.1 Accedo’s virtual reality video on demand prototype ............................................................ 25
  4.2 Heuristic evaluation procedure ............................................................................................. 27
  4.3 Interview and usability test procedure ................................................................................... 28
  4.4 Survey procedure .................................................................................................................. 32

5 Result and analysis .................................................................................................................... 34
  5.1 Heuristic evaluation result and analysis ................................................................................. 35
  5.2 Interview result and analysis .................................................................................................. 36
  5.3 Usability test result and analysis ............................................................................................ 38
  5.4 Survey result and analysis ...................................................................................................... 44
  5.5 Comprehensive analysis ........................................................................................................ 47
    5.5.1 Identifying the users’ goals, needs, behaviours, preferences, and context of use .......... 47
    5.5.2 Analysis of the discovered user experience problems .................................................... 48

6 Suggested guidelines .................................................................................................................. 51

7 Conclusion and discussion ......................................................................................................... 56
  7.1 Contributions and implications ............................................................................................... 56
  7.2 Methodological discussion ..................................................................................................... 57
  7.3 Future research ....................................................................................................................... 58
  7.4 Concluding remarks .............................................................................................................. 58

References .....................................................................................................................................

Appendix A: Heuristic evaluation inspection questions ..............................................................

Appendix B: Interview questions ................................................................................................

Appendix C: Questions after usability test ....................................................................................

Appendix D: Survey design ...........................................................................................................

Appendix E: Heuristic evaluation result .......................................................................................
# List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D</td>
<td>Two-dimensional</td>
</tr>
<tr>
<td>2,5D</td>
<td>Two and a half-dimensional</td>
</tr>
<tr>
<td>3D</td>
<td>Three-dimensional</td>
</tr>
<tr>
<td>3DUI</td>
<td>Three-dimensional user interface</td>
</tr>
<tr>
<td>AR</td>
<td>Augmented reality</td>
</tr>
<tr>
<td>DOF</td>
<td>Degrees of freedom</td>
</tr>
<tr>
<td>FOR</td>
<td>Field of regard</td>
</tr>
<tr>
<td>FOV</td>
<td>Field of view</td>
</tr>
<tr>
<td>GUI</td>
<td>Graphical user interface</td>
</tr>
<tr>
<td>HMD</td>
<td>Head-mounted display</td>
</tr>
<tr>
<td>UI</td>
<td>User interface</td>
</tr>
<tr>
<td>UX</td>
<td>User experience</td>
</tr>
<tr>
<td>VE</td>
<td>Virtual environment</td>
</tr>
<tr>
<td>VoD</td>
<td>Video on demand</td>
</tr>
<tr>
<td>VR</td>
<td>Virtual reality</td>
</tr>
<tr>
<td>VRGUI</td>
<td>Virtual reality graphical user interface</td>
</tr>
</tbody>
</table>
1 Introduction

Virtual reality (VR) was predicted to become the big technology trend of 2016 (Cellan-Jones, 2016). After almost 50 years in development where the technology has struggled to make the breakthrough (Williams, 2015), Perry (2016) claims that ‘The door to mass-market virtual reality is about to burst open.’ During the year, a variety of head-mounted displays (HMDs) for VR experiences – including Oculus Rift, Sony’s PlayStation VR, and the HTC Vive – will be released (Perry, 2016; Cellan-Jones, 2016).

Considering most HMDs has barely been launched, the published research addressing general usability and how these HMDs affect users is limited (Serge & Moss, 2015). User-centred research is, therefore, needed. Furthermore, O’Connell (2012) claim that designing VR for a good user experience (UX) requires new ways of thinking about usability principles.

When more commercial HMDs enter the market, the numbers of VR applications are expected to increase (Dorabjee, Bown, Sarkar & Tomitsch, 2015). One main factor for whether an application will succeed or not is the user interface (UI) (Molina, González, Lozano, Montero & López-Jaquero, 2003). While decades of research have laid the foundation for the technology of HMDs, there has been a lack of focus on the UIs (Dorabjee et al., 2015). Bowman (2013) claims that the principles of good three-dimensional (3D) UI design are more important than ever to understand. Compared to guidelines for graphical user interfaces (GUI), the design principles for 3D UI design are not nearly as well developed.

The development of VR GUI (VRGUI) presents unique challenges (Oculus VR, 2015a). One design principle for 3D interaction is to understand the interaction techniques available (Bowman, 2013). There is a new world of interaction, and the rules and guidelines are changing (Norman, 2010). System control for 3D UIs is in its infancy as a research topic, and there is a lack of empirical evidence for the usability of various system control techniques (Bowman, Kruijff, LaViola & Poupyrev, 2004). Don Norman, one of the most influential voices within the user experience industry, recently said ‘As we start moving into virtual reality (...) it is not at all obvious how you control things’ (Axbom & Royal-Lawson, 2016). No traditional input method is ideal for VR (Oculus VR, 2015b). Innovation and research are needed.

In VR, fundamental forms of interaction are movement, selection, manipulation, and scaling (Mine, 1995). One of the main techniques used in VR is direct user interaction, which can be defined as the use of gestures and movements of head and hands to specify interaction parameters. However, the most intuitive way to interact with a VRGUI according to Oculus VR (2015a) is through gaze tracking. Eye gaze direction can be approximated using the orientation of a user’s head (Mine, 1995). If a cursor is placed in the centre of the direction the user is currently facing, it can provide the user with the experience of controlling the system with his or her gaze direction (Oculus VR, 2015b). To the best of the author’s knowledge, this is the only interaction and input method that is currently compatible with all HMDs and VR sets used with smartphones without the need for any external input device.

One of the most widely researched fields, within the field of human-computer interaction, has been the use of eye gaze as an input modality to control systems (Bazrafkan, Kar & Costache, 2015). Even though researchers have put much effort in the field, the usability of eye gaze as an input modality remains marginal (Bazrafkan et al., 2015). Equally important is, that to the best of the author’s knowledge, the use of eye gaze direction approximated by head orientation as an input modality has barely been researched. Furthermore, there is a lack of a commonly used term for the interaction method and therefore, the suggested term ‘head gaze’ is used throughout this thesis.
With the introduction of Oculus Rift, the gaming industry became interested in VR (Dorabjee et al., 2015). However, today VR is becoming integrated across many different domains, such as medical, training simulations, and education. Mark Zuckerberg recently shared his vision for the future by saying ‘When you think about virtual reality, a lot of people first think about gaming. But I actually think video is going to be more engaging in a lot of ways’ (Zuckerberg, 2015).

Yu, Zheng, Zhao and Zheng (2006) state that the next major step in the evolution of media content delivery is streaming video on demand (VoD) over the Internet. VoD services allow users to take part of any video content at any time and place (Mustillo, Belanger, Fayad & Menzies, 1997). However, from the user perspective, using a VoD service, is not about accessing many different movies, it is about having a quiet evening with a spouse or indulging in a TV show (Goodman, Kuniavsky & Moed 2012). It is all about the experience.

Accedo is a company that develops TV application solutions with the mission to deliver attractive and successful experiences across a variety of connected video devices (Accedo, n.d.). They served as a research partner during the thesis and their developed VoD service prototype for VR - controlled by head orientation input - was tested in one part of the research.

The aim of this thesis is to contribute to the understanding of what aspects should be considered when designing VRGUIs - controlled by head orientation input - for qualitative UX. One contribution from the research is the definition of the users’ pragmatic and hedonic goals when using a VoD service in VR. Furthermore, the users’ expectations, needs, behaviours, preferences and predicted context of use were clarified. In addition, current UX problems of the prototype were discovered and analysed. However, the main contribution is seven unique developed UX guidelines. They are based on the research result and analysis in combination with recommendations discovered in the background literature review. Nevertheless, these are considered as a research foundation since they are not validated, and more research is required to confirm the relevance of them.

While this chapter serves as an introduction to the thesis, Chapter 2 defines and discusses the different core concepts of VR, VRGUIs, HMDs, head orientation input and head gaze, UX, and VoD services. Furthermore, it presents existing design principles related to the research and the research aim, an objective, expected contributions, and limitations. Chapter 3 describes the research design and the four methods that were used including a heuristic evaluation, interviews, usability tests, and a survey. The procedure of the used methods is presented in Chapter 4 together with a description of Accedo’s VoD service prototype for VR, which was used for the heuristic evaluation and the usability tests. Chapter 5 present the result and analysis of the research methods separately and comprehensively. The developed guidelines are presented and described in Chapter 6. Finally, Chapter 7 covers conclusions, contributions, a methodological discussion, and future research.
2 Background

This chapter aims to define and inform about the core concepts in the thesis and present previous related research. As a result of the recent commercialization of head-mounted displays and virtual reality for consumers, a number of the sources used in the literature review are based on websites, blogs, and video content. There is a shortage of published and peer-reviewed research within the field.

The first section of the background chapter defines virtual reality and describes synonyms and related concepts. Section 2.2 defines virtual reality graphical user interfaces and explain important factors and concepts related to it. The third Section presents the head-mounted display and its characteristics and disadvantages. Section 2.4 focuses on head orientation input and the suggested term head gaze for the interaction method. Thereafter, the term user experience is explained and in Section 2.5. Section 2.6 covers and describes video on demand services. In Section 2.7, existing design guidelines for 3D interaction and virtual reality graphical user interfaces are described. The final section, Section 8, defines the research question, aim, objective, expected contributions, and limitations.

2.1 Virtual reality and related concepts

Many different definitions of virtual reality (VR) have been used over time. In 1993, Adam (1993) defined VR as ‘a combination of various interface technologies that enables a user to intuitively interact with an immersive and dynamic computer-generated environment’. A couple of years later, the definition was simplified and described VR as an interactive, three-dimensional computer-synthesized environment (Pratt, Zyda & Keller, 1995; Barfield & Furness, 1995). Brooks (1999) decided to focus on the immersion in the definition that claimed VR experiences to be any in which the user is effectively immersed in a responsive virtual world. Years later, VR was seen as a medium in Sherman and Craig’s (2003) definition. They stated that VR is ‘a medium composed of interactive computer simulations that sense the participant’s position and actions and replace or augment the feedback to one or more senses, giving the feeling of being mentally immersed or present in the simulation (a virtual world).’ At the same time Bowman, Kruijiff, LaViola, and Poupyrev (2004) emphasized on how the VR is seen and controlled in their definition that VR is something seen from a first-person point of view that is also controlled by the user in real time.

In Sherman and Craig’s (2003) definition of a VR experience, they describe four key elements. These are a virtual world, sensory feedback as a response to the users’ input, interactivity, and immersion. A virtual world is the content of a given medium but can also be used as a synonym for VR. Essential to VR is the sensory feedback, which means that the system gives the users feedback based on their physical position. For the VR to seem authentic, it should be interactive and respond to the users actions. The last key element is physical immersion. Physical immersion is defined as the feeling of ‘bodily entering into a medium...’ (Sherman & Craig, 2003). This while immersion can be described as the feeling of being deeply engaged (Coomans & Timmermans, 1997). The concept of ‘being immersed’ is often described by media, and generally refers to a feeling of being involved in an experience (Sherman & Craig, 2003). The VR community is also using the term presence to present the same concept. Users’ sense of copresence and understanding of the environment are affected by the level of immersion they are experiencing (Bowman, Coquillart, Froehlich, Hirose, Kitamura, Kiyokawa & Stuerzlinger, 2008).

VR has many synonyms and also relate to several other concepts. The term virtual reality was originally coined by Jaron Lanier (Lanier & Biocca, 1992; Machover & Tice, 1994; Krueger, 1991;
Before the expression became popular, earlier experimenters, such as the VR pioneer Myron Krueger, used the term artificial reality (Machover & Tice, 1994; Biocca & Levy, 1995; Wexelblat, 1995). Krueger (1991) defined artificial reality as a reality that ‘...perceives a participant’s action in terms of the body’s relationship to a graphic world and generates responses that maintain the illusion that his or her actions are taking place within that world’. Nowadays artificial reality coincides with what is generally referred to as VR (Sherman & Craig, 2003). Historical definitions often described VR in the terms of technological hardware but Steuer (1992) discussed a need for a new definition based on the concepts of presence and telepresence to move past that. As previously described, presence can be used synonymously with immersion and telepresence can be defined as ‘the ability to directly interact (...) with a physical real, remote environment from the first-person point of view...’ (Sherman & Craig, 2003). Telepresence is also a term that can be synonymously used with VR (Ashline & Lai, 1995).

Another commonly used term is virtual environments (VE) (Sherman & Craig, 2003). Barfield and Furness (1995) define VE as ‘...the representation of a computer model or database which can be interactively experienced and manipulated by the virtual environment participant(s)’. According to Sherman and Craig (2003) there are two definitions of VE, one describing it as a virtual world and the other describing it as ‘...an instance of a virtual world presented in an interactive medium such as virtual reality’. VE is also often used synonymously with VR (Bowman et al., 2004; Earnshaw, Gigante & Jones, 1993; Sherman & Craig, 2003). Some researchers prefer to use the term VE over VR, since the hype over VR have led to unrealistic expectations being associated to it (Bowman et al., 2004; Earnshaw et al., 1993).

Cyberspace is another synonymously used term for VR (Ashline & Lai, 1995). William Gibson coined the term in his science fiction novel Neuromancer that was published 1984 (Machover & Tice, 1994). If the term is not used synonymously with VR, cyberspace can be defined as ‘a location that exists only in the minds of the participants, often as a result of technology that enables geographically distant people to interactively communicate’ (Sherman & Craig, 2003).

Related to VR is augmented reality (AR). AR enhances the real world by computer-generated overlays (Adam, 1993; Sherman and Craig, 2003). Zhou, Dun, and Billinghurst (2008) define AR as a technology, which allows computer-generated virtual imagery to exactly overlay physical objects in

---

**Figure 1. An overview of the key concepts, synonyms, and related terms to virtual reality described throughout Section 2.1.**
real time. While some definitions are dependent on certain technology, Azuma (1997) state that AR is any system that combines real and virtual, is interactive in real time and is registered in three dimensions. Similarly to AR is mixed reality (Bowman et al., 2004). It can be defined as a continuum including both VR and AR.

To conclude, various definitions of VR often have several similarities but focus on different components. Adam (1993) describes VR in words of technology, and Brooks (1999) focus on the immersive feeling while Sherman and Craig (2003) defines VR as a medium. The definition used for this thesis is that VR is an interactive, three-dimensional computer-synthesized environment (Pratt et al., 1995; Barfield & Furness, 1995) seen from the first-person point of view and controlled in real time by the user (Bowman et al., 2004). Furthermore, it is extended to include Sherman and Craig’s (2003) four key elements. Figure 1 provides an overview of the key concepts, synonyms, and related concepts to VR described throughout the section.

Next section will define and focus on graphical user interfaces in VR. It will present different content display techniques, challenges, and solutions.

2.2 Virtual reality graphical user interface

The user interface (UI) is the part of a computer and its software that people can see, listen to, talk to or in any other way understand and direct (Galitz, 2007). In over 20 years, the graphical UI (GUI) has dominated the way people interact with computers (Harper, Rodden, Rogers & Sellen, 2008; van Dam, 1997). Furthermore, it has been a standard since the 1990s when Apple first introduced Macintosh (Butow, 2007).

The standard components of GUIs are called widgets (Hartson & Pyla, 2012). They include for example windows, buttons, menus, toolbars, and scrollbars (Butow, 2007). Based on widgets, and commonly associated with GUIs, is the acronym WIMP, which stands for windows, icons, menus, and pointing devices (Molina et al., 2003). Many GUIs are also designed following the desktop metaphor. The metaphor was originally developed to make the interaction easier for users when the first personal computer arrived (Hartson & Pyla, 2012). It is created based on the users’ existing knowledge about their own desktop environments and the objects used around them and applies the same concept with the same functionality in the GUI. The metaphor is an example of how the interaction can be simplified with cognitive affordance (defined in Section 2.5).

A 3D UI (3DUI) is defined as a UI that involves 3D-interaction (Bowman et al., 2008). Research within 3DUIs has traditionally been closely connected to VR. An important part of 3DUIs is depth information, which helps users to interact with the application. Depth can be visualised in various ways with the help of different kind of cues. These can be divided into four categories (Bowman et al., 2004). The first category is monocular, static depth cues. They are depth information that can be viewed from a static image by a single eye (Bowman et al, 2004). Moreover, they are also referred to as pictorial cues and include occlusion, linear and aerial perspective, shadows and lighting, texture, gradients, relative size and height relative to the horizon. The second category is oculomotor cues. They derive from muscular tension in the viewer’s visual system, described as accommodation and convergence (Bowman et al, 2004). Thirdly motion parallax is what causes objects close to the viewer to move more quickly across the visual field than objects further away from the viewer. The last category is binocular disparity and stereopsis. Binocular disparity refers to the difference between the two images that each eye sees at the same time. This while stereopsis is the effect that occurs when these two images fuse into one single stereoscopic image providing a depth cue (Bowman et al, 2004).
One common approach to denote hierarchy in a GUI is with the use of size, contrast, and colour (Sundström, 2015). While these tools are still available in VR, they are used differently. For example, size is based on the distance between the user and the piece of content. Furthermore, content can be displayed in various ways. Three different content display techniques are the ‘head-up display’, keeping the content locked to the environment on an object or floating free (see Figure 2). The ‘heads-up display’ locks the content at a set distance from the viewer. This display technique follows the user if s/he decides to turn around or move in the virtual environment, keeping the content constantly facing towards the user at the set distance. When the content is locked to the environment, the content stays in the same location when the user moves, allowing the user to explore it from different angles. Moreover, the content can be locked to an object or float free (Sundström, 2015).

Sundström (2015) describe how designers are trying to force two-dimensional (2D) solutions into a 3D space even though they have a full field of vision to design in. However, he claims one reason to use 2D in 3D is to facilitate for the user’s small cone of focus (further described in Section 2.3) when they are using a VR set, a type of head-mounted display, to experience VR. Furthermore, a solution to the small cone of focus is to use a common tile menu. Variations in tile menus include if they are flat, curved or surrounding the user (see Figure 3). A flat tile menu looks like a wall to the user. This solution makes it difficult to read text or see images in perspective. The curved tile menu places the tiles faced towards the user, making it easier to read text and view images. However, according to Sundström (2015), the best solution is to create a surrounded tile menu where the menu is placed around the user in the virtual environment. Further recommendations include keeping the menu content limited and placing content of less importance pushed out of immediate view, and the small cone of focus, but still accessible.
The next section will describe the technology used to experience VR. The head-mounted display is defined and important characteristics are further described, including possibilities and limitations of the technology.

2.3 Head-mounted display

A VR system consists of a display, a tracking device, a computer image generator, a three-dimensional database, and application software (Adam, 1993). Moreover, different displays are available to suit different tasks and experiences. The displays can, for example, be CAVE-like (cave automatic virtual environment) surround projectors, panoramic displays, workbench projectors, or desktop displays (Brooks, 1999). However, the display developed for and providing the greatest sense of immersion is the head-mounted display (HMD) (Coomans & Timmermans, 1997; Adam, 1993). HMDs are wearable devices in the form of goggles (Dorabjee, et al., 2015) that blocks out the vision of the physical world and projects computer-generated stereoscopic images onto two display screens close to the user’s eyes (Kavakli & Jayaratna, 2005). Advances in hardware technology have recently led to the production of HMDs for consumers, such as the Oculus Rift, which is suitable for the use of immersive VR applications including gaming, simulation, and film (Carnegie & Rhee, 2015).

Various terms are used for the technology. Oculus Rift is referred to as an HMD (Serge & Moss, 2015) while Google Cardboard and Samsung Gear VR are two examples of so-called VR sets (Sundström, 2015). While Google (2016) and Sundström (2015) refer to these as VR sets, other authors, such as Lamkin (2016), use the term VR headsets for both HMDs and VR sets. Furthermore, Oculus VR (2015a) refers to Gear VR as a VR headset and not a VR set. The virtual reality society (2016) claim that the terms VR headsets and VR glasses are simply synonyms to HMDs and that if the HMD attach straight to the users head and present visuals directly to the users eyes it could be considered an HMD in the broadest sense. Therefore, this thesis will use the term HMD for all variations including VR sets, VR headsets, and VR glasses (see Figure 4 for three examples of different HMDS).

Different HMDs uses power from different technologies. While Oculus Rift plugs into a computer, Sony PlayStation VR needs to be connected with a PlayStation (Lamkin, 2016). Furthermore, Samsung Gear VR and Google Cardboard use smartphones as their processors and displays. The smartphones’ gyroscopic sensors and positioning systems are used to accurately track the users’ head movements.

When describing visual display devices, such as HMDs, important characteristics include the field of regard (FOR) and field of view (FOV) (Bowman et al., 2004). FOR refers to the amount of the physical space surrounding the user. It can also be explained as the maximum number of degrees of visual angle that can be seen instantaneously on a display. It is measured in degrees of the visual angle, so if a cylindrical display were built where the user could stand in the middle, the display would have a 360-degree horizontal FOR. The FOV must be less than or equal to the maximum FOV of the human

![Figure 4. Examples of three different head-mounted displays. Oculus Rift developer kit 2 to the left, Google Cardboard used with a smartphone in the middle and Homido used with a smartphone to the right.](image)
visual system, which is approximately 200 degrees. HMDs usually have a small FOV but a large FOR. A lower FOV may decrease immersion and results in ‘tunnel vision’ while higher FOV can decrease the resolution and introduce distortion (Bowman, Datey, Ryu, Farooq & Vasnaik, 2002).

In HMDs, the FOV is about 94 degrees (Oculus VR, 2015c). However, users can change the orientation of their heads to see more of their surroundings. Chu (2014) describes how the head orientation can be changed comfortably and topmost. People can rotate their heads to the left about 30 degrees in a comfortable manner. Topmost, they can rotate their heads 55 degrees to the left. The same degrees apply to when people are rotating their heads to the right as well. Furthermore, people can comfortably turn their heads up, while 60 degrees is the topmost. Moreover, people can turn their heads with the chin down. While 12 degrees is comfortable, the topmost is 40 degrees. These degrees, suggested by Chu (2014) are visualised by Mike Alger (2015a) (see Figure 5). When these degrees are combined with the FOV in an HMD, it creates 77 degrees to the right and to the left of the centre where the users can see objects comfortably and about 102 degrees, to the right and to the left of the centre, where the users can see objects in a more strained way (Alger, 2015a).

Another important aspect of how objects should be placed in virtual environments experienced with HMDs concern the sense of depth. Chu (2014) claims that the stereoscopic effect will be lost if objects are placed further than 20 metres away from the user. More importantly, if objects are placed too close to the user, it will cause cross-eyes and double vision. To avoid straining the users’ eyes, objects should not be placed closer than a distance of 75 centimetres from the user (Oculus VR, 2015b). Mike Alger (2015a) has used these recommendations, but another recommendation considering the minimum distance claiming it to be 50 centimetres, and designed figures illustrating where the objects can be seen comfortably and meaningfully and where he thinks the main content should be displayed (see Figure 6).
However, there is a difference between general content and a GUI. Lubos et al. (2014) suggest that 3D selection tasks that require high precision should be placed close to the users' eyes. Moreover, Oculus VR (2015a) states that the ideal distance to place a UI in VR is between 1 to 3 meters away from the user.

Other important characteristics of visual display devices include the spatial resolution (Bowman et al., 2004). The spatial resolution is a measure of visual quality related to pixel sizes. It is often measured as dots per inch but is also affected by the screen size and the distance between the user and the visual display. Inside an HMD as for example the Google Cardboard or Gear VR a single screen of a smartphone is split into two, one for each eye, dividing the resolution (Sundström, 2015). Moreover, the user focuses on the centre of this area and creates a cone-of-focus that quickly fall-off towards blurriness. This leads to a small low-resolution area to design for (Sundström, 2015).

Further characteristics of visual display devices include screen geometry, light transfer mechanism, refresh rate, and ergonomics (Bowman et al., 2004). The screen geometry is related to the shape of the screen. Visual displays can be rectangular, L-shaped, hemispherical, or hybrids. Moreover, the light transfer mechanism decides how the light is transferred onto the display surface (Bowman et al., 2004). Light can be front projected, rear projected, or as a laser light directly projected onto the retina. This is of importance since different types of 3DUI interaction techniques are applicable for different light transfer methods. The refresh rate describes in which speed the visual display device refreshes the displayed image. This is usually measured in hertz as refreshes per second. Finally, ergonomics cover how comfortable the visual display device is to use. Bowman et al. (2004) emphasize that comfort is especially important for visual display devices such as HMDs.

Typical disadvantages of HMDs include system latency, limited FOV, peripheral vision, and ergonomic discomfort (Bowman et al., 2008). When displaying VR with the use of an HMD, one of the most concerning aspects is the common experience of simulator sickness (SS) (Serge & Moss, 2015). It can cause physical reactions such as headaches, nausea, dizziness, and eyestrain (Carnegie & Rhee, 2015). This usually generates negative experiences and could lead to confusion or even fainting (Moss & Muth, 2011).

In the following section, input devices and interaction methods for VR experienced with HMDs are further explained. Moreover, it focuses on the head orientation input and the suggested term head gaze.

### 2.4 Head orientation input and head gaze

The UI is composed of input and output (Galitz, 2007). Input is the technique that people use to communicate their needs and desired actions to the computer while the output is how the computer mediates the result of the users’ directives. An interaction method can be defined as how a person uses an input or output device to perform a task in a human-computer dialog (Foley, Van Dam, Feiner & Hughes, 1990). Three main techniques that are used in VR are physical controls, virtual controls, and direct user interaction. Physical controls are the use of for example buttons, sliders, and steering wheels while virtual controls are control devices presented as virtual objects. Direct user interaction can be defined as the use of gestures and movements of head and hands to specify interaction parameters (Mine, 1995). Fundamental forms of interaction in VR include movements, selections, manipulations, and scaling (Mine, 1995).

One concern with interaction in VR is that the real-world metaphor does not allow efficient large-scale placement of objects in the virtual environment (Bowman & Hodges, 1997). Techniques developed to solve this and grab as well as manipulate remote objects are for example arm-extension techniques and ray-casting techniques. One example of an arm-extending technique is the
Go-Go interaction technique (Poupyrev, Billinghurst, Weghorst & Ichikawa, 1996). Ray-casting is an interaction technique that uses a virtual light ray, specified by the user’s hand, to point and grab objects (Mine, 1995).

Another concern is that interaction in 3D mid-air is physically demanding and often can hinder user satisfaction and performance (Chan, Kao, Chen, Lee, Hsu & Hung, 2010). VR selection problems can often be efficiently solved with pointing techniques (Lubos, Bruder & Steinicke, 2014). However, they require a level of abstraction that is potentially less natural compared to the virtual hand metaphor (Lubos, Bruder & Steinicke, 2014). Nevertheless, direct interaction leads to higher performance than manipulation of objects at a distance (Mine, Brooks & Sequin, 1997).

In computer graphics systems, motion trackers can be used for five primary purposes (Welch & Foxlin, 2002). The first use is for the users view control, which means that it provides the users with position and orientation control and simulates a first-person viewpoint for people using HMDs. Secondly, motion trackers can be used for navigation. A third way to use motion trackers is for object selection and manipulation. Finally, motion tracking can be used for instrument tracking and avatar animation. Tracking errors can contribute to destroying the sense of perceptual stability, causing sickness or degrading task performance (Welch & Foxlin, 2002).

Oculus VR (2015b) claims that gaze tracking is one of the most intuitive ways to interact with a VRGUI. The orientation of the users’ heads can be used to approximate the users’ gaze direction (Mine, 1995; Mine et al., 1997). This can be simulated to the users by placing a cursor floating in the middle of their FOVs (Mine, 1995). The cursor is then centred in the direction the users are currently facing, giving the users the experience of controlling the system with their gaze direction. This creates an intuitive way to select an item by simply looking at it (Mine et al., 1997). Users can interact with a VRGUI through this method as if it was a mouse or a touch device (Oculus VR, 2015b).

To make a selection, the users ‘look’, but actually turn their heads, towards an object. The selection is then indicated with a standard selection signal. Dwell time is a technique that can be used to select one object among several displayed (Jacob, 1991; Lankford, 2000). This means that when a user continues to look at an object for a certain predetermined time, the object will be selected. A long dwell time will decrease the amount of unintended selections that are made, but it will also reduce the responsiveness of the UI. Jacob (1991) claims that if the result of selecting the wrong object can easily be undone, then a shorter dwell time can be used. For this approach, a dwell time of 150-250 milliseconds can be used with good results. This provides the users with the subjective feeling of a highly responsive system. A longer dwell time, over ¾ second, is never useful because it is not natural for eyes to fixate on one spot for that long and the user can also suspect that the system has crashed. A good use of the object selection interaction is to request further details about an object. The details should be displayed on a separate area where such information is always displayed in the UI (Jacob, 1991).

The method of using head orientation input to estimate the users’ gaze direction and place a cursor in the centre of the users’ FOV is referred to as ray-casting by Oculus VR (2015b). However, this stands in contrast to the previously described definition in this section by Mine (1995) who claims that ray-casting is the use of a virtual light ray. With regard to the conflicting definitions, the term ray-casting was not used as a name for the method in this thesis. Mine (1996) uses the name ‘look-at menus interaction technique’ for the technique, but he also includes the use of an external input device in the interaction. Furthermore, Mine’s suggested name has not been used or published by other researchers and therefore, will not be used in this thesis either. To the best of the author’s knowledge, there is no commonly used term for this interaction method and therefore, the suggested term ‘head gaze’ will be used throughout this thesis.
Key issues in eye tracking in VR involve the Midas touch problem and gaze-point tracker inaccuracies leading to incorrect actions and selection (Bazrafkan et al., 2015). The Midas touch problem relates to peoples’ eye movements that are frequently unintentional, which makes it problematic to use them as an input modality to control a system (Jacob, 1991). To develop a useful eye-tracking interface, the Midas touch problem needs to be avoided (Jacob & Karn, 2003). Another problem with limiting the users’ eye movements is that it will also limit the users’ spatial degrees of freedom (DOF) (Carnegie & Rhee, 2015). However, when using the head gaze interaction, the movement will occur through head orientation rather than eye movements, which compensates for this drawback.

A challenge with using gaze is that the user is always looking at the displayed content but may not always look at the VRGUi (Oculus VR, 2015a). Because of this, the user can ‘lose’ the UI and not realize it is still available, which can result in confusion when the application does not handle the input the way the user expects. This problem can be solved by making the UI close when the user goes outside some FOV or by automatically dragging the interface with the view as the user turns. A third solution is to place an icon somewhere on the periphery of the screen that indicates that the user is watching the UI and then allow this icon always to track with the users view (Oculus VR, 2015a).

A considerable amount of previous research has focused on eye gaze as an augmented input in combination with another input device or technique (see for example Fono & Vertegaal, 2005, Miniotas, Špakov, Tugoy & MacKenzie, 2006, Yamato, Monden, Matsumoto, Inoue & Torii, 2000, Salvucci & Anderson, 2000 and Zhai, Morimoto & Ihde, 1999). This research is not presented in this thesis since this thesis focus on eye gaze direction estimated by head orientation used as the only input and interaction technique. Likewise, research describing eye gaze estimation where it clearly requires eye-tracking technology, and the research is not believed to be applicable for the head gaze interaction method, has also been excluded.

The next section will define user experience and explain important factors influencing it. Furthermore, it will also describe how the user experience can be measured.

### 2.5 User experience

User experience (UX) is associated with several different meanings (Forlizzi & Battarbee, 2004). Nielsen and Norman (n.d.) summarize their definition of UX by stating that “‘User experience’ encompasses all aspects of the end user’s interaction with the company, its services, and its products.” The International Organization for Standardization (ISO) define UX, with their standard 9241-210, as users perception and response that has resulted from the use or the anticipated use of a product, a system, or a service (ISO, 2010). Moreover, UX includes the users believe, emotions, perception, preferences, physical, and psychological response and also behaviours and performance that happens before, during, and after the use. Additionally, there are three important factors influencing the UX. They are the system, the user, and the context of use (ISO, 2010). A well-known definition, highlighting those three important factors is the one described by Hassenzahl and Tractinsky (2006):

> UX is a consequence of a user’s internal state (predispositions, expectations, needs, motivation, mood, etc.), the characteristics of the designed system (e.g. complexity, purpose, usability, functionality, etc.) and the context (or the environment) within which the interaction occurs (e.g. organisational/social setting, meaningfulness of the activity, voluntariness of use, etc.).
Hassenzahl (2008) also describes two different dimensions of UX. Pragmatic quality is what he refers to as ‘...the product’s perceived ability to support the achievement of “do-goals”, such as “making a telephone call”...’ while hedonic quality refers to ‘...the product’s perceived ability to support the achievement of “be-goals”, such as “being competent”...’ The pragmatic quality is related to the products utility and usability in relation to a potential task while the hedonic quality focus on the Self, like the question of why someone own and use a particular product. Hassenzahl claims that the pragmatic quality might lead to the fulfilment of hedonic quality while the hedonic quality contributes directly to the UX. While pragmatic attributes emphasize the individuals’ behavioural goal, the hedonic attributes emphasize the individuals’ psychological wellbeing (Hassenzahl, 2008).

A key component to assuring qualitative UX is the usability (Hartson & Pyla, 2012). According to the ISO standard 9231-11, usability is defined as ‘the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use.’ (ISO, 1998).

However, efficiency and effectiveness are not enough to make users satisfied, since they are also looking for emotional satisfaction (Shih & Liu, 2007), which is another important part of the UX (Hartson & Pyla, 2012). The emotional impact is the component of the UX that influences users feelings. This includes effects such as pleasure, fun, the joy of use, aesthetics, desirability, pleasure, novelty, originality, sensation, coolness, engagement, novelty, and appeal. This can also include deeper emotional factors such as self-expression and self-identity, the pride of ownership and a feeling of contribution to the world. The importance of this is expressed by Hassenzahl, Beu and Burmester (2001) who claim that ‘The most basic reason for considering the joy of use is the humanistic view that enjoyment is fundamental to life.’

One concept used to improve the UX of systems, and a technique that supports users during interactions are affordances (Hartson, 2003). There are four different types of affordances including cognitive, physical, sensory, and functional affordances. The first type is cognitive affordances that help users with their cognitive actions, such as thinking, learning, and remembering. This while physical affordances help users performs physical actions, such as clicking, touching, pointing, and moving objects. Moreover, sensory affordances aim at helping users with sensory actions, such as seeing, hearing, and feeling things. The fourth and final type are functional affordances that help users get things done and use a system to do work (Hartson, 2003).

This thesis will use Hassenzahl and Tractinsky’s (2006) definition that UX is a consequence of a user’s internal state, the characteristics of the designed system, and the context within which the interaction occurs. Furthermore, it is extended also to include the pragmatic and hedonic quality (Hassenzahl, 2008) where usability is an important component (Hartson & Pyla, 2012). Finally, the definition also emphasises the emotional impact. More importantly is that UX cannot be designed; it can only be experienced. Qualitative UX is therefore considered to be UX fulfilling the users’ subjective hedonic and pragmatic goals with respect to the context of use with high usability and a desirable emotional impact.

In the following section, the application of video on demand services is briefly described. The section defines the service and important factors of it.

### 2.6 Video on demand service

Video on demand (VoD) allows people to access videos, for example movies, from video servers on a broadband network (Li, Liao, Qiu & Wong, 1996). Users can control their experiences and choose which video content to consume and at what time and location they prefer to consume it (Mustillo et
al., 1997). Moreover, the service is often personalized and involves one-to-one interaction between the user and the service (Choi, Reaz & Mukherjee, 2012).

To develop VoD services that lead to qualitative UX, it is highly important to understand the users' behaviour while interacting with the service. From the user perspective, Netflix media streaming service is not about accessing many different movies, it is about having a quiet evening with a spouse, making time pass at an airport or indulging in a TV show (Goodman et al., 2012).

The design of VoD services usually depends on the request patterns of the users (Choi et al., 2012). An example is that videos popularity correlates with the request rate. A well-known fact is that video popularity follows the Zipf distribution, which is an equation that shows request frequency of certain movie titles. Moreover, videos popularity may decrease with time since the interest decreases after people have watched it. The popularity can also change due to the introduction of new videos or new recommendations (Yu et al., 2006).

Another important factor in VoD services is the daily access pattern of the users’ behaviour (Choi et al., 2012). After a video is started and the user starts to watch it, the user might need to interact with the service to change the state of the video. Possible user interactions can be classified as play, resume, stop, abort, pause, jump forward, jump backward, fast search and slow motion (Li et al., 1996). Out of these interactions, pause is the most commonly used (Costa, Cunha, Borges, Ramos, Rocha, Almeida & Ribeiro-Neto, 2004). For long videos, jump forward and jump backward are equally common actions. However, the video playtime is usually quite short (Yu et al., 2006). More than half of all started videos are played less than 10 minutes and 70% are terminated within the first 20 minutes.

Since many VoD services, like Netflix, offer thousands of options of content, a central problem of VoD services is that the experience has become more about selecting a piece of content than simply watching it (Vanhemert, 2014). Users are forced to choose content before the watching experience can begin. Vanhemert claims that ‘Netflix is great when you want to watch something, but it’s terrible when you want to watch anything.’

The next section will briefly describe existing guidelines related to the design of VRGUIs. It summarizes earlier research and guidelines that practitioners currently use.

### 2.7 Research related design principles for virtual reality graphical user interfaces

This section presents existing guidelines of relevance for the research aim. The guidelines include how to design for 3D interaction, VR, VRGUI, head gaze, and video in VR. Table 1 shows an overview of the 10 design guidelines discovered and presented in this section.

**The first discovered guideline** is to allow users to be lazy (Wilson, 2011). Bowman (2013) claims that ‘Users in HMDs don’t want to turn their heads, much less move their bodies.’ If the users are forced to stretch their necks and turn in their seat, it could leave them sore, which degrades the UX (Sundström, 2015). A recommendation is, therefore, to not require users to move their head or body more than necessary. Bowman (2013) similarly emphasises not to neglect user comfort. Large movements could also be a problem since it would increase the likelihood of the user clashing with the surrounding environment outside of the HMD (Sundström, 2015).
The second discovered guideline, relating to the first, is to allow ‘magic’ interaction (Bowman et al., 2004). Interaction in 3D is often thought of as ‘natural’, but for many novice users operating a 3D UI is anything but natural (Bowman, 2013). Bowman et al. (2004) recommend violating assumptions about the real world in the virtual world. If the interaction is designed to be ‘magical’ instead of ‘natural’, it allows users to go beyond the perceptual, cognitive, and physical limitations of the real world (Bowman et al., 2004). Instead of the user having to walk up to an object and reach out with the hand to grab it, s/he can use a virtual laser pointer to grab it from a distance.

The third discovered guideline is that the number of degrees of freedom (DOF) the user is required to control, in general, should be as few as possible (Mine et al., 1997). When designing for 3D interactions, it is important to understand the design space and the interaction techniques that can be used (Bowman, 2013). The design space of 3D interaction varies depending on the number of devices and mappings available. Bowman et al. (2008) state that 2D tasks are cognitively easier than 3D tasks. Furthermore, real-world tasks, at least most of them, are not fully 3D (Bowman et al., 2008). As a result, most humans are used to dealing with 2D or two and a half-dimensional (2,5D) and do not have the skills to deal with 3D problems. Bowman et al. (2008) therefore, suggest creating 2D alternatives for as many tasks as possible in 3DUIs to increase the usability.

Furthermore, it is recommended to use 2D input devices over 3D input devices (Bowman et al., 2008). A comparison of input specifications between mouse- or pen-based systems and 3D technologies reveals that 2D technologies are much more precise (Teather & Stuerzlinger, 2008). When an input with high DOF is used for a task that requires a lower number of DOFs, the result can be that the task is unnecessarily difficult to perform (Bowman, 2013). An example is when users want to select a menu item, which is inherently a one-dimensional task, and they need to position their virtual hands within a menu item to select it. The UI then requires too much effort from the user to perform the task. Moreover, forcing users to position their hands in three-dimensions to press a virtual button slows down the interaction and increases user frustration (Mine et al., 1997). This relates to Bowman’s (2013) recommendation to keep the UI simple. He claims that if the user’s goal is simple, there should be a simple and effortless technique to reach it.
The guideline can be followed in several different ways. Keeping the number of DOFs to a minimum can be accomplished by using a lower-DOF input device, ignoring some of the input DOFs, or by using physical or virtual constraints (Bowman, 2013).

**The fourth guideline** discovered is that VR needs affordances to indicate which objects the user can interact with and when the interaction takes place (Sundström, 2015). One way to help users perform the correct action is to use real-world metaphors (Bowman et al., 2004; Sundström, 2015). Moreover, Alger (2015b) suggests that designers should use the human instincts to their advantage when they are designing VRGUIs.

**The fifth guideline** discovered is to design for the hardware (Bowman, 2013). Bowman claims that what might work for one display or with one device rarely work exactly the same way on a different system. When using the same UI on a different display or device, the UI, and the interaction techniques often need to be modified. This is called the migration issue (Bowman, 2013). Another important part of why the design should be for the technology is to have the ergonomic and physical constraints of the technology in mind (Chen, 2014). If this is not considered, interactions are more likely to be disrupted.

**The sixth guideline** for easy-to-use interaction techniques is to not use floating, interpenetrating or invisible objects (Bowman et al., 2008). Bowman claims that floating objects should be the exception and not the rule. He refers to the real world, which the user is familiar to, where floating objects barely exist and most objects are attached to other objects. Interpenetrating objects should be avoided since many novice users have problems recovering from such situations. Furthermore, a recommendation is never to force users to interact with invisible objects (Bowman et al., 2008).

**The seventh guideline** is to display the VRGUI at a reasonable and comfortable distance from the user (Oculus VR, 2015a; Chen, 2014). Lubos et al. (2014) suggest that 3D selection tasks that require high precision should be placed close to the users eyes. Oculus VR (2015a) claim that the ideal distance is usually between one to three meters away from the user. This since placing an object too close to the user can cause a jarring experience. As previously described in Section 2.3, the recommendation is to keep objects at a distance of at least 75 centimetres from the user (Oculus VR, 2015b) but not further than 20 metres away (Chu, 2014).

**Guideline eight** is to use a gaze cursor or crosshair when the gaze is used as the input modality (Oculus VR, 2015a). Using gaze selection without a cursor or crosshair has been reported as more difficult to use since the user has to estimate where the head gaze is currently located.

**The ninth guideline** is to design the video screen in VR to cover less than 70 degrees of the horizontal FOV (Oculus VR, 2015a). Oculus VR claims this is the size that is required to allow users to view the full screen without turning their heads.

**The tenth and final guideline** is to avoid rapid movements and not to drop frames or loose head tracking (Sundström, 2015; Wilson, 2011). If frames are dropped or the head tracking is lost it will appear as an error in the users’ consciousness (Wilson, 2011). This could cause the user to feel sick. Also, rapid movements could have the same effect (Sundström, 2015). They can also cause the user to feel disoriented in the virtual environment.

The following section will describe the research aim and objective. It will also cover the expected result and limitations of the research.
2.8 Research aim and objective

Perry (2016) claims that ‘The door to mass-market virtual reality is about to burst open.’ Considering most HMDs has barely been launched, the published research addressing general usability and how these HMDs affect users is limited (Serge & Moss, 2015). Furthermore, three decades of research has laid the foundation for the technology of HMDs, but there has been a lack of focus on the UsIs (Dorabjee et al., 2015). Moreover, Oculus VR (2015b) claims that no traditional input method is ideal for VR and that innovation and research are needed. One of the most researched fields of human-computer interaction has been the use of eye gaze as an input modality (Bazrafkan et al., 2015). However, the usability of eye gaze as input still remains marginal. Moreover, there is a shortage of research concerning the use of head orientation input to estimate gaze direction.

The aim of this thesis is to contribute to the understanding of what aspects are important to consider when designing virtual reality graphical user interfaces - controlled by head orientation input - for qualitative user experiences. In order to achieve this aim, the following research question has been developed:

What aspects should be considered when designing a graphical user interface in virtual reality - controlled by head orientation input - for a qualitative user experience?

The user interface (UI) is the part of a computer that people can see, listen to, talk to or in any other way understand and direct (Galitz, 2007).

Virtual reality (VR) is defined to be an interactive, three-dimensional computer-synthesized environment (Pratt et al., 1995; Barfield & Furness, 1995) seen from the first-person point of view and controlled in real time by the user (Bowman et al., 2004). Moreover, it includes a virtual world, sensory feedback, interactivity, and immersion (Sherman & Craig, 2003).

Head orientation input refers to the way that the orientation of the user’s heads can approximate the user’s gaze direction and be used as an input modality to control a UI. One way to simulate this to the user is by having a cursor floating in the middle of their field of view (Mine, 1995). The interaction method that uses head orientation input to approximate the user’s gaze direction is in this thesis referred to as head gaze.

User experience (UX) can be defined as a consequence of a user’s internal state, the characteristics of the designed system, and the context within which the interaction occurs (Hassenzahl & Tractinsky, 2006). Furthermore, it is also extended to include the pragmatic and hedonic quality (Hassenzahl, 2008) where usability is an important component (Hartson & Pyla, 2012). Finally, the definition also emphasises the emotional impact. However, UX cannot be designed; it can only be experienced. Qualitative UX is therefore considered to be UX fulfilling the users’ subjective hedonic and pragmatic goals with respect to the context of use with high usability and a desirable emotional impact.

In order to achieve the thesis aim, the following objective has been created:

Investigate and analyse the concepts involved in the research question and how they are affecting the user experience.

The achievement of the aim of this research enables the possibility to develop guidelines for the design of graphical user interfaces for virtual reality – controlled by head orientation input – for a qualitative user experience.
2.8.1 Expected contributions

Several contributions are expected from the research. The main expected contribution is guidelines that should be taken into consideration when designing VRGUIs – controlled by head orientation input – for qualitative UX. Furthermore, the research is expected to contribute to the understanding of VRGUIs should be designed for the application of VoD services – controlled by head orientation input – for qualitative UX. Except for this specific application, the research is also expected to contribute to the understanding of the core concepts individually. This could include the understanding of how VRGUIs should be designed for other applications than VoD services or how to design UIs controlled with head orientation input in other contexts. Additionally, Accedo will also benefit based on the fact that their existing VRGUI prototype will be tested and evaluated during the research. Finally, the research is expected to contribute to a foundation for future research within the field.

2.8.2 Limitations

The research aim is to contribute to the understanding of what aspects are important to consider when designing VRGUIs for qualitative UX. This is limited to design as visual appearance and interaction, and not to design as a process of developing VRGUIs. The VRGUIs in focus are also limited to the ones experienced with the use of HMDs.

Also the input of head orientation is limited to the interaction method of head gaze as the only interaction method used to control the system with no additional input device or interaction method.

The application of the research also serves as a natural limitation, since the UX will only be researched for VoD services. Moreover, the users and target group for the research are based on Accedo’s previously used expectations that the users are between 20 to 45 years old. They have also chosen only to focus on users that are not wearing glasses since all the HMDs do not yet support users with glasses. This target group was further limited to people that were experienced users of VoD services but with no or limited experience of VR. Furthermore, the target group was limited to people that use technology on a daily basis without being novice users or technology experts. Finally, the target group was limited to people that could comfortably speak and read Swedish and English since the usability tests and interviews were performed in Swedish but Accedo’s VR prototype that was tested is developed in English.

The research is also limited to the application of VoD services containing only 2D video content. With regard to the research being located in Sweden, it will mainly show a result of users part of the Swedish culture and not cover cultural differences. Furthermore, the author’s knowledge of how to measure degrees of FOR and FOV in VR is limited and can only be roughly estimated. The final limitation is that the background research does not cover all existing guidelines for VRGUIs or head orientation input.
3 Method

While the previous chapter presented the background and defined the different concepts involved in the research question and aim of the thesis, this chapter defines and describes the methods used for the research. The first section focuses on the research design while the following section continues to further describe each of the methods used. Section 3.2 defines the heuristic evaluation, and the third section focuses on interviews while Section 3.4 explains usability tests. The final section defines the method of using surveys for research purposes.

3.1 Research design

In order to answer the research question, the goal of the research was to collect data about what influence the user experience (UX) when head orientation input and head gaze is used to control a graphical user interface (GUI) in virtual reality (VR). Furthermore, the application video on demand (VoD) services was also considered an important concept in the research.

However, there is a shortage of VoD services in VR controlled by head orientation input and therefore, the research could not exclusively focus on testing different applications. Since the research question involved several concepts for this specific application, there were several different combinations of these concepts together or alone that could be researched to gain insights.

Qualitative research focuses on gaining context and insight regarding user behaviour (Unger & Chandler, 2012). Equally important is that it provides results that are closer to the core of the users’ subjective experiences compared to quantitative research approaches (Anderson et al., 2010). Therefore, qualitative research was selected as the main research approach for the thesis. Additionally, a limited amount of quantitative data was also collected and analysed together with the qualitative data.

User research can be used to better understand the users and test their behaviour in a specific domain (Unger & Chandler, 2012). It is a process with the aim to investigate how people perceive and use services and products (Goodman, Kuniavsky & Moed, 2012; Anderson, McRee & Wilson, 2010). The key concepts of user research are to understand the users needs on both an intellectual and intuitive level and see things from their perspective (Anderson et al., 2010). One important part in user research and UX design is the concept of goals since products are only seen as tools to accomplish higher-level goals for users.

The first step in performing user research is to define the primary user groups (Unger & Chandler, 2012). A method to develop the first definition of the user groups is to create a provisional definition based on knowledge available in the project team. The next step is to plan for the users participation in the research by the use of different methods. At the beginning of a project, this could be interviews, surveys, focus groups or contextual inquiry’s (Unger & Chandler, 2012).

There are several reasons to use more than only one method for the research. Using two or more research methods provides a richer picture of the user than one method can provide on its own (Unger & Chandler, 2012). Moreover, triangulation is the process of using several methods to find problems in a system (Wilson, 2014). The method is used to compare separate sources of findings to discover consistencies and discrepancies (Barnum, 2011). Additionally, it is the most powerful method to determine which problems are the real problems in a product (Wilson, 2006). The problems that are repeatedly found are more likely to be the real problems. Another reason to use the method is to eliminate the risk that the found problems are found based on the specific method.
used (Banister, Dunn, Burman, Daniels, Duckett, Goodley, Lawthom, Parker, Runswick-Cole, Sixsmith, Smailes, Tindall & Whelan, 2011). Finally, triangulation leads to a more contextual level of the collected information (Biggerstaff, 2012) and enriches the analysis of findings (Barnum, 2011). Time and financial resources are usually what decides the amount of methods used in the same research (Barnum, 2011).

For UX evaluation, Hartson & Pyla (2012) recommends to improvise, adapt, and to customize methods and techniques to fit the evaluation goals and needs. This could implicate excluding steps, introducing new steps, and changing the details of a step.

The methods that were selected for this research are heuristic evaluations, interviews, usability tests and a survey. They were selected with the aim to collect UX data about the different involved concepts. Table 2 show how the different methods gather insights about different concepts.

Alternative methods that could have been used were for example focus groups (Unger & Chandler, 2012) or perspective-based UI inspection, cognitive walkthrough, or a formal usability inspection (Wilson, 2014). Time constraints and the authors existing knowledge of the other methods made these methods prioritized lower when the selection was made. In the following sections of the chapter, each method will be defined and explained further including the motivation behind why they were selected for the research.

### 3.2 Heuristic evaluation

Heuristic evaluations are a type of user interface or usability inspection where an individual or a team evaluates a prototype or a product based on a list of heuristics (Nielsen, 1993; Nielsen & Molich, 1990). The main goal of using the method is to find as many usability and design problems as possible with limited economic and time resources (Wilson, 2014).

One advantage of heuristic evaluations is that it is an easy and intuitive method to learn (Hartson & Pyla, 2012). The evaluations are relatively time effective to perform, requires no special resources or real users, and can provide additional breadth and complement other assessment techniques (Wilson, 2014). Therefore, the method is efficient to use in the beginning of a UX process (Hartson & Pyla, 2012). When the method is used before usability testing, it can contribute to identifying goals for the testing (Barnum, 2011). Additionally, the use of the evaluation before the usability test has the advantage that the facilitator of the usability test will be familiar with the UI and have a basic understanding of the problems that the users might experience during the test.

On the contrary, there is a debate whether the heuristic evaluation is a good method to find the ‘real’ problems (Wilson, 2014). One example is that a problem for a beginner might be perceived as a positive feature for an expert and vice versa. The quality of the heuristic evaluations results also varies based on the experience of the evaluators. Hartson & Pyla (2012) describe that there will always be certain UX problems that appear in real live user-based interactions that will not appear in any heuristic evaluation or other inspection method. Another risk is that heuristic evaluation, like other inspection methods, creates false negatives, which are situations where the inspectors
identified problems that turned out to not be a real problem or not as important as the evaluator thought (Hartson & Pyla, 2012). However, these possible risks are perceived as minimised with the use of complementing research methods and the triangulation of the results between them.

Several steps are included in the planning of a heuristic evaluation (Wilson, 2014). The first step is to recruit the evaluating team. Width and depth of the functions that are expected to be covered by the evaluation together with the evaluators’ experience, domain knowledge, and the complexity of the UI determine the size of the team (Wilson, 2014). Most heuristic evaluations are performed by a team of typically two or three inspectors but one is always better than none and sometimes a single evaluator is enough for early project stages (Hartson & Pyla, 2012).

The next step of the planning is to decide the list of heuristics that should be used for the evaluation (Wilson, 2014). Nielsen and Molich (1995) developed a set of heuristics that are often mentioned (Barnum, 2011; Hartson & Pyla, 2012) and commonly used (Unger & Chandler, 2012). Nielsen (1995) later decided to refine the heuristics (see Table 3). Other heuristics are for example Gerhardt-Powals (1996) ten research-based heuristics and Weinschenk and Barkers (2000) 20 general heuristics. Unger and Chandler (2012) recommend keeping the list of heuristics to a manageable size since a too large amount of them could make the evaluation unnecessary hard to perform.

Next Hartson & Pyla (2012) recommends evaluators to create inspection questions based on the chosen heuristics. The following part of the planning is to create a method for how the problems should be collected, organized, and presented (Wilson, 2014). One way is to use a chart, either printed or digital, with different fields where the inspectors can document the problem, the related heuristic, the severity of the problem, and where the problem was found.

Table 3. Nielsen’s 10 usability heuristics (Nielsen, 1995).

<table>
<thead>
<tr>
<th>Nielsen’s 10 usability heuristics (Nielsen, 1995)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Visibility of system status</td>
</tr>
<tr>
<td>The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.</td>
</tr>
<tr>
<td>2. Match between system and the real world</td>
</tr>
<tr>
<td>The system should speak the users’ language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.</td>
</tr>
<tr>
<td>3. User control and freedom</td>
</tr>
<tr>
<td>Users often choose system functions by mistake and will need a clearly marked ‘emergency exit’ to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.</td>
</tr>
<tr>
<td>4. Consistency and standards</td>
</tr>
<tr>
<td>Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.</td>
</tr>
<tr>
<td>5. Error prevention</td>
</tr>
<tr>
<td>Even better than good error messages is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.</td>
</tr>
<tr>
<td>6. Recognition rather than recall</td>
</tr>
<tr>
<td>Minimize the user’s memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.</td>
</tr>
<tr>
<td>7. Flexibility and efficiency of use</td>
</tr>
<tr>
<td>Accelerators -- unseen by the novice user -- may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.</td>
</tr>
<tr>
<td>8. Aesthetic and minimalist design</td>
</tr>
<tr>
<td>Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.</td>
</tr>
<tr>
<td>9. Help users recognise diagnose and recover from errors</td>
</tr>
<tr>
<td>Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.</td>
</tr>
<tr>
<td>10. Help and documentation</td>
</tr>
<tr>
<td>Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user’s task, list concrete steps to be carried out, and not be too large.</td>
</tr>
</tbody>
</table>
practice the method (Wilson, 2014). Materials and information, such as user profiles and environments where the product will be used need to be prepared. Furthermore, access to the product needs to be prepared. The final step of the preparations is to decide which approach the evaluators should follow. One approach is that the evaluators follow a list of task scenarios and another alternative is to follow a list of goals to be reached during the evaluation (Wilson, 2014).

The evaluations should be performed individually (Wilson, 2014). First, the evaluators need to read the background information including goals, personas, and environment descriptions. Next, they need to inspect the task scenarios. During the evaluation, each problem should be listed individually. This simplifies the process of identifying if the problems are local or global. After all the individual evaluations are performed, a summarizing list of all the found problems should be created. The list should be sorted in a way that is suitable for the project. It could be sorted based on the severity of the problems or grouped after the different parts of the system (Wilson, 2014).

3.3 Interview

User interviews are individual conversations with users from the primary target group (Unger & Chandler, 2012). Goodman et al. (2012) describe that interviews are needed to really understand users’ experiences. It is a method appropriate to gather data about the users honest opinions, preferences, and attitudes (Unger & Chandler, 2012).

When using interviews, it is sometimes difficult to collect straightforward opinions and collect information about attitudes and context (Unger & Chandler, 2012). The participants will not always say what they believe and sometimes they will answer to a different question than the one being asked (Goodman et al., 2012). This can usually be solved by observing the user during the interview and look for cues in behaviour such as hesitation or body language. Another solution is to show interest in the participants’ opinions and being persistent asking the same question again with slight changes in wording and emphasis (Goodman et al., 2012).

Recruiting is the process of finding, inviting, and scheduling the right people to use in the research (Goodman et al., 2012). The process consists of three basic steps: determining the target audience, finding members of that audience, and convincing them to participate in the research.

The process of determining the target audience starts with writing about the product’s target audience and then by deciding if there is a specific part of the target group that will result in better feedback for the specific research (Goodman et al., 2012). It is also important to consider the amount of previous experience they have with the product, with competing products, and if there are characteristics that should be avoided. When the recruiting starts, a recommendation is to try to avoid recruiting participants with the exact same qualities if there are variations allowed within the target group (Goodman et al., 2012). An example is if the target group spans over ages from 20 to 30 years old people, then the participants should be of different ages in between and not all 25 years old.

Tools for recruiting include for example online forums, mailing lists, social networking services, asking acquaintances, or using ads (Goodman et al., 2012). Friends and family should be avoided since their relationship to the interviewer will affect their behaviour. When deciding a number of participants to recruit, it is also important to plan for recruited participants that cancel or do not show up (Goodman et al., 2012).

The preparations before the interviews include creating a list of questions that will research the users’ previous experience of the system, their goals, and needs (Unger & Chandler, 2012). It is also
important to decide how structured the interviews should be. A semi-structured approach can be used if the amount of data is more important than consistency. The semi-structured approach starts with the interviewer asking prepared questions but allows the conversation to follow a natural pattern to continue researching interesting commentaries from the participants (Unger & Chandler, 2012).

Goodman et al. (2012) recommend planning the interview guide to be followed to start with simpler questions and then continue with more serious questions. Furthermore, the questions should be open and flexible so the participants can introduce topics and questions that the interviewer would not have been able to predict (Hartson & Pyla, 2012). If a participant is asked to choose from several options, s/he will choose one even if none of the options suit the participant’s opinion (Goodman et al., 2012). Moreover, the questions should not cover several subjects or be binary. When the interview is completely planned, it should be tested to control if any of the questions are hard to understand or if the time schedule needs to be adjusted (Unger & Chandler, 2012).

Before the interview starts, the participants need to be informed of their role in the project and the terms for their participation (Vetenskapsrådet, 1990). All of the participants should be offered as much confidentially as possible and be asked to leave consent for their participation.

During the interview, it is important to make the participant feel comfortable talking to the interviewer since it will help the person to answer questions in an honest way (Goodman et al., 2012). A technique to make the participant feel comfortable is to show interest for topics non-related to the interview since it will help the person feel more relaxed. Furthermore, the interviewer needs to work with an open mind to not exclude topics that seem too trivial. In order to relieve the interviewer’s cognitive abilities during the interviews, a sound recorder can be used. Then the interviewer can keep his or her focus on the participant and the questions instead of focusing on taking notes (Goodman et al., 2012).

3.4 Usability test

Usability tests are based on users trying to perform typical tasks in a system while a moderator observes and in some case ask questions to understand the users’ behaviour (Unger & Chandler, 2012). The tests are suited to evaluate existing systems or prototypes that need to be improved. Formative testing is when a product is in development, and the goal is to find and rectify problems (Barnum, 2011). These are typically based on small studies repeated during development. The resulting data often take the form of problem descriptions and design recommendations (Sauro & Lewis, 2012).

Usability tests are commonly used in the design phase of a project when a design is created, tested and then improved and tested again (Unger & Chandler, 2012). Barnum (2011) recommends combining usability tests with heuristic evaluations since each of the methods brings value to the process of creating products that are useful, usable, and desirable. The combined result of the two methods will also provide a stronger confirmation of the problems that are the ‘real’ problems (Barnum, 2011).

The first step of planning a usability test is to establish test goals, decide the product that should be tested, including where, and how the test should be performed (Barnum, 2011; Unger & Chandler, 2012). Barnum (2011) describe that the test goals should explain the part of the users’ experience that is of focus to research in the test. When the product and the test goals are established, the next step is to plan where the tests should be located. Usability tests can be done in labs, conference rooms, or in the field. Furthermore, the location also includes preparing the right equipment for the
test (Hartson & Pyla, 2012). If quantitative UX data is planned to be collected it is useful to have timers or software prepared (Vermeeren et al., 2002). However, it is also possible to collect high-precision timing data from video recordings of the session. Usability tests can be performed for benchmarking to establish metrics of the product or as a comparison of designs (Barnum, 2011). The typical test, often performed in formative usability testing, is to give the users a number of tasks within scenarios and gather feedback on their experience of the product.

Usability tests can be used as a quantitative or qualitative method (Unger & Chandler, 2012). The advantages of using the test as a quantitative method is that it generates results that can be controlled statistically at later iterations and it lowers the likelihood that an individual moderator has affected the result. The advantages of using the usability test as a qualitative method is that is usually less expensive since fewer users are needed for the test. It is also a more available method for moderators with limited experience of performing the test and leads to data with a higher information density (Unger & Chandler, 2012).

According to Nielsen and Landauer (1993), the number of usability problems that will be found in a usability test in relation to the amounts of users participating can be calculated using a mathematical model. While Nielsen and Landauer’s study shows that at least 15 users need to be in the test to discover all the usability problems, Nielsen (2000) claim that the best result comes from testing no more than five users per test and running as many small tests as can be afforded. A usability test with five users will lead to about 85% of the usability problems being found (Nielsen, 2000). Recommendations for recruiting of the participants follow the same recommendations as previously described for interviews in Section 3.3.

Essential elements that need to be incorporated for good results in small usability studies include defining a user profile, creating task-based scenarios, using a think-aloud process, and making changes and testing again (Barnum, 2010). The user profile is the basis for recruiting participants for the study and is previously covered in Section 3.3. Task-based scenarios are tasks embedded within scenarios. They usually include realistic descriptions based on users’ goals. Using a think-aloud process means encouraging participants to share their thoughts verbally with the moderator during the test. An advantage of the method is that it will become easier for the moderator to understand why the users do what they do without the moderator having to guess what they are thinking (Barnum, 2011).

Before the tests, a participant incentive should be determined and prepared (Barnum, 2011). The incentive should be experienced as a gift and not as a bribe. This in order to not influence the remarks the participants make. Appropriate incentives include cash, gift cards, or debit cards.

During the usability test, the first step is to meet and greet the participant (Barnum, 2011). The following part is to give the user a pre-test briefing. In order to always share the same information with all the participants, the moderator should use a script for this part of the test (Barnum, 2011). After the participant is verbally informed, s/he should be asked to sign a consent form containing the same information (Hartson & Pyla, 2012). Vetenskapsrådet (1990) strongly recommends following four specific requirements in the research. The first requirement is to inform the participant of the purpose and aim of the research. Secondly, the participant should always leave consent for the participation and that they are aware that the participation is voluntary. The third requirement is that the participants should be offered as much confidentially as possible and that their private information should be protected from others. Finally, the participants should be informed of how the gathered data will be used (Vetenskapsrådet, 1990).

The final step before the test starts is to inform the participants about the think-out-loud process (Barnum, 2011). A typical test usually lasts between 30 minutes to four hours for an individual
participant (Hartson & Pyla, 2012). However, Hartson and Pyla recommends keeping the session under two hours, since it will become more difficult for the participant to stay focused if the test is longer.

During the test, different types of data can be gathered. Sauro and Lewis (2012) recommend not to let the sample size stop a test from using statistics to quantify the data and inform design decision. Completion rates, for example, are the most fundamental of usability metrics (Nielsen, 2001). Also task times, the amount of errors, and satisfaction rates can be used as measurements (Sauro & Lewis, 2012). Moreover, problems can also be quantified in terms of frequency and severity.

3.5 Survey

Surveys are series of questions used to find patterns in behaviour and preference among larger numbers of people (Unger & Chandler, 2012). They are useful when the aim is to research preferences rather than performance. Surveys also serve as good supplements to other forms of research. Sauro and Lewis (2012) claim that surveys are one of the easiest methods to collect attitudinal data. Furthermore, many participants do not always like to state negative opinions out loud but are often willing to express them when faced with a ranking system (Unger & Chandler, 2012). Another advantage is that surveys are valuable to supplement objective UX evaluations data with subjective data directly from the users (Hartson & Pyla, 2012). Additionally, the method is easy to use for both analyst and participant and does not require a specific location.

Surveys can yield both quantitative data as well as qualitative user opinions (Hartson & Pyla, 2012). Typically they contain a combination of open-ended comments, binary responses and Likert-type rating scales (Sauro & Lewis, 2012).

When planning a survey, it is important to consider the users targeted and which method for distributing the survey will give the best result (Unger & Chandler, 2012). For example, if the target users tend to be active Internet users, sending them an online survey could be the best solution.

Another important factor when creating a survey is to take into consideration the amount of time the participants are willing to spend answering it (Unger & Chandler, 2012). If compensation is offered to the participants, the survey can generally be longer, but if no compensation is offered, the survey need to be kept short to help ensure people will complete it. Unger and Chandler recommend keeping the completion time required between five to ten minutes. Another part of the planning is to decide how long the survey will run. Alternatives include letting the survey run until a certain amount of responses is collected or until a certain deadline, depending on which has the highest priority (Unger & Chandler, 2012). As previously described for interviews and usability tests in Section 3.3 and Section 3.4 it is also important to inform the respondents of the purpose and aim of the research (Vetenskapsrådet, 1990).

Finally, it is important to consider which tool will be used to collect and analyse the data (Unger & Chandler, 2012). If the survey is running online, the tool may have options for viewing and analysing the result. Otherwise, a method for how to enter the data into another analysing tool is needed.
4 Procedure

This chapter aims to describe how the different methods were executed during the research. The first section will describe the prototype that was used in the heuristic evaluation and the usability tests. Section 4.2 explains the procedure of the heuristic evaluation. The following section describes how the interviews and the usability tests were performed. Finally, Section 4.4 describes how the survey was executed.

4.1 Accedo’s virtual reality video on demand prototype

As previously stated, there is a shortage of available video on demand (VoD) services in virtual reality (VR) controlled with head orientation input and head gaze. However, the company Accedo had a high fidelity prototype of exactly that kind of application. The prototype is developed with the project name Himalaya, and it has been presented publicly but not officially usability tested. With respect to the aim of the research, the prototype has been used as a product to be usability tested in order to gather UX data and insights.

The prototype is available in the form of an application under development. It can be downloaded to any smartphone using Android or iOS with Accedo’s permission. The prototype can then be explored using any head-mounted display (HMD), or so-called VR set, that is used with a smartphone. Resolution of the smartphone screen and the quality of the HMD are two factors influencing the quality of the experience.

When entering the application, a big video screen is displayed in the centre of the users’ field of view (FOV) and depending on how the users’ change the orientation of their heads, different functions and menus will appear (see Figure 7). The different information areas will appear as animations when the users’ change the orientation of their heads. Sundström (2015) previously (see Section 2.2) described different display methods in VR and this prototype’s content is displayed floating in the air but is locked to the world. Furthermore, it is displayed as a curved tile menu, directed towards the user.

A crosshair is constantly displayed in the centre of the user’s FOV and when the user turns his or her head, the crosshair follows and remains at a constant position in the centre of his or her FOV. The crosshair is used as a cursor and when the users place the crosshair over a UI element that the user can interact with it will display hovering information about it and start a circular progress bar counting down the time until the selection.

Figure 7. How the different menu’s appear when the user turns his or her head in Accedo’s prototype.
equal as a mouse click, will occur. The predetermined time before the click starts after the head gaze is placed over a UI element is 200 milliseconds and the time it takes before the selection is performed is 2000 milliseconds. Different feedback follows from different UI element selections.

If a user places the head gaze cursor straight forward at the screen s/he will see the video screen and if the head gaze is retained at the same location for a predetermined period of time, the screen will zoom in and become larger to the user. This while the surroundings will turn completely black so the only thing the user sees is the video screen. If the user changes the orientation of his or her head slightly down, a menu will open up over the lower part of the screen (see the centre image in Figure 8). The menu contains functions to play, pause, fast-forward, rewind, and rewind 10 seconds. This menu will be referred to as the play menu.

If the user locates the head gaze over the arrows pointing to the left, the video will start rewinding. First, it will follow a speed of two times the normal running time but later it will accelerate step by step until it reaches a speed of 16 times the original speed if the user retains the head gaze over the arrows. The fast-forwarding arrows pointing to the right use the same functionality but reversed. The rest of the icons, except the orange button in the centre will trigger the circular progress bar to start

---

**Figure 8.** An overview over all the information areas and menus that open when the user changes his or her head orientation in Accedo’s virtual reality prototype of a video on demand service.
and if the user retains the head gaze over the icon until the progress bar is full, the selection will be completed. The orange icon in the centre is not possible to interact with and is just a visual detail.

If the user turns his or her head to the left of the screen, a menu of round buttons will appear (see the left image in Figure 8). This menu is referred to as the main menu and it contains buttons for the functions of different content categories. The different buttons can be chosen to display movies, TV shows, channels, or to exit the VR experience. After a selection is completed, the user should turn his or her head down below the screen to continue looking for a specific video in the display menu.

The display menu encompasses video posters and titles of different videos in a 3D carousel menu (see the image at the bottom in Figure 8). The menu changes content based on which selection is chosen in the main menu. In the menu, the user can scroll to the right and left side to access more content. The scrolling will begin when the user places the head gaze cursor on the left or right side of the displayed menu content. Moreover, the scrolling will continue until the user moves the head gaze cursor to the centre of the displayed content. A progress bar below the menu shows how much content is available and where among the content, the user is currently placed. A selection of one video is performed if the user places the head gaze cursor over it and stays over it until the full selection time is completed. Then the video will start to play above the display menu on the video screen.

If the user turns his or her head up, the settings menu will appear (see the top image in Figure 8). It contains round buttons with different on and off functionality. The user can choose to turn on a leak light function displaying a background light behind the screen or turn on and off sound effects for buttons and interactions. S/he can also change the shape of the screen to be flat or curved and turn on and off scenery of Himalaya Mountains that appear behind the screen and the menus on the black background.

If the user turns his or her head to the right side past the screen edge, information about the video currently playing will appear (see the image on the right side in Figure 8). The information is only for display and the user can not interact with it more than read it. If a movie or TV show is currently running on the screen the area contains the video title, year of release, a short summary of the videos story, the director of the video, actors in the video, and also the video poster. If a channel is currently running on the screen, the information area will show the name of the channel and a describing text about it.

The video content in the prototype is 2D and only a few trailers of several combined videos are currently available as placeholders in the application. When a user selects and starts a specific video, the video content displayed will always be one of the trailers.

4.2 Heuristic evaluation procedure

The first method used in the research was the heuristic evaluation. Based on time constraints and with the approach of using the method quick and informal it was performed by the author alone. No recruiting of evaluators was therefore required. Nielsen’s 10 usability heuristics (1995) (previously presented in Table 2 in Section 3.2) were selected for the evaluation based on the author’s previous knowledge and the recommendations from Unger and Chandler (2012) of keeping the list of heuristics to a manageable size to not overcomplicate the procedure of performing it.

Based on Hartson and Pyla’s (2012) recommendations, inspection questions were created for the evaluation. The first questions was ‘Is it clear to the user, what is happening in the system?’.
questions per heuristic was created and used during the evaluation (see Appendix A for the complete list of questions).

The remaining part of the planning and performing of the evaluation followed Wilson’s (2014) recommendations (see Section 3.2). The problems were documented in a digital chart with fields for the problem description, the problem-related heuristic, the severity of the problem, and where in the system the problem was found. Since the author was already experienced with the method, there was no need to practise the method before performing the evaluation. A list of tasks, based on assumed important actions the user would want to perform, was created for the evaluation. The list included starting a video, pausing a video, fast-forwarding and rewinding in a video, changing to another video, and finding a specific video.

Finally, the access to the product was prepared and previous documentation was inspected. Documentation included documents from the development of the prototype with thoughts concerning the development, the target group and information about competitors’ solutions.

The evaluation was performed by the author answering the inspection questions while trying to perform the list of tasks previously created. All the problems were listed separately. When the evaluation was completed, the list was sorted based on the severity of the found problems and problems that appeared to be global but was listed two or more times as local were merged into one global problem instead.

4.3 Interview and usability test procedure

The interviews and the usability tests were combined and performed after each other during the same session with the same participants. This was done with respect to the time constraints and limited resources of the research but also with the aim to reach a higher quality of the gathered data. Users were first interviewed and then participated in the usability test. The interview served as a natural introduction for the participants to start thinking about their VoD preferences and behaviour, and what they were expecting from a video in VR. Furthermore, the data from the interviews and the usability tests could be compared for each participant, with the aim to reach a better understanding of the users and the user experience (UX) of the prototype.

Following the recommendations of Goodman et al. (2012) (see Section 3.3), the first step was determining the target audience. The target audience for the Himalaya demo prototype developed by Accedo was found in one of Accedo’s documents describing the audience as 20 to 45 year old people without glasses since many HMDs do not work with glasses yet. Furthermore, the group was described as comfortable speaking English, since English is the language used in the prototype. They were also described as comfortable watching online videos, such as the ones in VoD services, and not just linear TV. The part of the target group that was selected as the one that would give the best feedback considered previous experience the users might have with a product or competing products and if there were characteristics that should be avoided. Based on this, the group was limited to people with no earlier or very limited experience of VR to capture the first impression, experience, and behaviour. The group was also limited to users who understood and could speak both English and Swedish since the interview was performed in Swedish in Sweden. Furthermore, the group was limited to experienced VoD service users. They were defined as people using computers and smartphones on daily bases without being beginners with technology or technology experts. The limitations defined the user profile for the recruiting process to begin, also following Barnum’s (2011) recommendations for usability tests.
Following Nielsen’s (2000) recommendations, the amount of users to be used in the research was decided to be five. An additional user was added in order to plan for possible fallouts, following recommendations from Goodman et al. (2012). The recruiting also followed recommendations from Goodman et al. (2012) (see Section 3.3). Facebook was used to recruit acquaintances and the author’s friends and family shared the post, which made it reach further than only the author’s social network. When selecting participants among interested people, variation among them was strived for.

Three men and three women of the ages between 27 and 45 years old were recruited. Following Barnum’s (2011) recommendations, the participants were offered gift certificates for the cinema as an incentive. Each participant was booked for a separate session scheduled to last 90 minutes. The time of the session was based on recommendations from Hartson and Pyla (2012) to keep each session under two hours.

The interview questions were created based on Unger and Chandler’s (2012) recommendations (see Section 3.3). With the aim to research the users’ previous experience, goals, and needs a semi-structured approach was selected. Following suggestions from Goodman et al. (2012), the interview guide started with simpler questions and then continued with more serious questions. Furthermore, the questions were created to be open and flexible, based on Hartson and Pyla’s (2012) recommendations.

The first part of the interview involved questions about the users’ previous experience and knowledge of VoD services. It was followed with questions researching their behaviour, goals, opinions and values regarding VoD services. One question was for example ‘Where do you usually use VoD services?’.

The second part of the interview covered the users previous experience of VR and their opinions and expectations of video in VR. One question was for example ‘How do you feel about watching video content in VR?’. Finally, the participants were asked about their previous experience and knowledge of using their head or eyes to interact with systems (see the complete list of questions asked during the interview in Appendix B).

The usability test was planned to follow recommendations from Barnum (2011) and Unger and Chandler (2012). The first step according to them is to establish test goals for the test. The goals were based on the result of the heuristic evaluation. Goals for the usability test were to test if the VRGUI was easy to learn and if the users ability to perform tasks improved as they became familiar with the system. Furthermore, the test was aimed to research if users could find the information they needed in order to complete a task without assistance. Another goal was to see if the users could effectively and successfully reach different goals that were assumed to be typical for VoD users. Moreover, another aspect was to research if the users experienced errors and how well they were able to recover from them. Finally, a test goal was to research the participants experience during the test and the different tasks included.

The next step was to decide a location for the testing. Conference rooms at the Accedo office in Stockholm were booked for the purpose. Following Hartson and Pyla’s (2012) recommendations, the equipment for the rooms was prepared as well.

A decision was made to mostly gather qualitative data but also a small amount quantitative data, such as completion rates and completion times. The prototype was installed on an iPhone 6 plus, since it was the smartphone with the largest screen available and since the author is previously most familiar working with iOS devices. An iPhone 6 plus screen has a 1920-by-1080-pixel resolution at 401 PPI (Apple, 2016). The HMD selection was also based on availability. Homido is an HMD that allows
users to enter VR with their smartphones (Homido, 2014). It has settings to set the eye to lens distance, custom-made VR lenses that offer a 100-degree field of view (FOV), interchangeable contact surface, interchangeable distance setting and is compatible with most recent smartphones (Homido, 2014).

To allow the moderator to focus on moderating the test and asking questions, following recommendations from Goodman et al. (2012), sound and video recording was used during the sessions. For the interview questions when the prototype was not being tested, an iPhone 5s was used for sound recording and when the prototype was tested, a system camera with recording functionality and a tripod was used. In order to see and hear what the participants were seeing and hearing during the usability test, the iPhone was connected to a TV screen in the room with an HDMI cable and an HDMI to lightning adapter. Usually the HMD is used with earphones connected to the smartphone with the application for the most immersive experience but since this would have prevented the moderator from hearing what the participant was hearing and also limit the possibility of giving instructions and asking questions to the participant during the test, the sound was led to the TV instead. This might have affected the experience but was a solution to a problem that might not have had any other easy solution. The full set of equipment and how everything was set up during the test can be seen in Figure 9.

Figure 9. The setup of equipment and how it was used during the usability test.

Table 4. List of task-based scenarios used in the usability test.

<table>
<thead>
<tr>
<th>Task-based scenarios used in the usability test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. This is your first time using this application so explore it and share spontaneous reactions and thoughts. Tell me what you think you can do here.</td>
</tr>
<tr>
<td>2. Imagine that you would like to know the name of the video that is currently running on the video screen. Try to find it (and remember that the video that is currently running on the video screen is only a placeholder).</td>
</tr>
<tr>
<td>3. Imagine that you would like to change to another video. Find and start another video than the one that is currently running on the screen.</td>
</tr>
<tr>
<td>4. Imagine that you would like to fast forward. Could you please do it?</td>
</tr>
<tr>
<td>5. Imagine you would like to pause the video that is currently running. Could you please pause it?</td>
</tr>
<tr>
<td>6. Imagine that you now would like to continue watching the video. Could you please start it again?</td>
</tr>
<tr>
<td>7. Imagine that you feel like watching the TV show ‘The Simpsons’. Could you find and start episode one of season one?</td>
</tr>
<tr>
<td>8. Imagine that you now feel like watching the TV show ‘Mad Men’ instead. Could you find it?</td>
</tr>
<tr>
<td>9. Now you are curious about all the functionality in the service so please explore it and try to find as many functionalities as possible. Please tell me how you think when you are looking for them and what you think you can do with them.</td>
</tr>
</tbody>
</table>
Barnum (2011) claims that task-based scenarios are required to have good results from a usability test. Therefore, task-based scenarios were developed, focusing on the test goals and the users’ goals. A total of nine task-based scenarios were prepared (see Table 4). Following recommendations from Vetenskapsrådet (1990) (see Section 3.3) a consent form was prepared for the usability test. The moderator script for the test was also prepared with reminders to explain the think-aloud method for the participants before the test started.

Finally, following recommendations from Barnum (2011) and Unger and Chandler (2012), the whole interview and usability test were tested during a pilot test. The pilot test was performed with the aim to configure questions and control the time required for the whole session. No changes were made after the pilot test since no problems were discovered.

Following the recommendations previously described in Section 3.4, the participant was first greeted and pre-test briefed by the moderator following a prepared script. All participants were asked to sign a consent form after they were informed of the purpose of the study, how the data would be used, that their participation was completely voluntary and that they were offered a certain level of confidentiality. The interview was sound recorded and when it was finished, the participant was informed about the usability test. They were told they would control the UI with their heads and approximated gaze but were not given details about how this would work.

The moderator started the video recording after confirming with the participants one last time and then started the application on the smartphone, placed it in the HMD and asked the participants to place it on their heads. Participants were informed about simulator sickness and reminded of the purpose of the test. They were also informed of the limitations of the prototype, such as the lack of real video content. Furthermore, they were informed about the procedure of the scenarios and the way the moderator would follow their actions looking at the TV screen connected to the iPhone in their HMDs. Finally, they were informed about the think-out-loud method.

The moderator started the video recording and then started the application, placed it in the HMD and asked the participants to put it on their heads. Before they entered the VR world, they were informed of the first task-based scenario in order to be able to start straight away and in order to capture first impressions. The first task-based scenario was to explore the application and share reactions and thoughts. They were asked to describe what they thought they could do in the application.

The second scenario was to figure out the title of the video that was currently running. The third scenario of the test was for the participant to find and start another video. After that, the participants were asked to fast forward in the video. The fifth and sixth scenario was to play and pause the video. Scenario seven described that they wanted to watch episode one of season one of the TV show The Simpsons. They were asked to find and start the episode. The following scenario asked them to find the TV show Mad Men. In the last scenario, the participants were encouraged to explore further and find as much functionality as possible.

After the test, participants were asked question about their experience and their answers were sound recorded (the full list of questions can be seen in Appendix C). One question was for example ‘Was there something that surprised you?’ and another was ‘Did it feel like you could control everything you wanted to control in the user interface?’. In total, most interview and test sessions were completed within 70 minutes but one of the sessions almost used the full scheduled time of 90 minutes. The session ended with the users receiving the incentive for the participation.
4.4 Survey procedure

The survey was carried out with the aim to research the users’ preferences and attitudes regarding VoD services. Since the aim of the study was wide, the target group was extended. The target group for the survey was decided to be people between the ages of 20 to 45 years old with experience of using VoD services. Their experience of VR was not relevant for the purpose of the survey and was therefore not included as a part of the target group definition.

The preparation of the survey followed Unger and Chandler’s (2012) recommendations (see Section 3.5). Since the aim was to collect as many responses as possible, and the time to gather data was limited and no incentive was available, the survey was limited to five questions, with a few sub-questions. An approximation and two pilot tests of the survey showed the time to complete the survey to be about five minutes. The pilot tests were successful and no problems occurred so no adjustment was made.

In the introductory part of the survey, the respondents were informed that only respondents within the target group were asked to answer it. The respondents were then informed about the survey’s purpose, how the collected information would be used and the conditions for their contribution, following Vetenskapsrådets (1990) recommendations for research. Finally, important terminology in the survey was defined.

The questions were of different types, some with rating scales for how frequently the user started a VoD service with a certain goal and another question was with ranking choices of what functionality that was the most important. Two questions were designed with a comment box for open answers and one was designed with multiple choices.

The first question was a statement of ‘This rarely/often do I start a video on demand service to...’ and then followed by five different statements that the respondents were asked to answer. The five different statements were:
1. Find something to watch when I do not know in advance what I would like to watch.
2. Find something to watch when I almost know in advance what I would like to watch.
3. Find a specific video when I know in advance what I would like to see.
4. Continue to watch a video I was watching previously.
5. Control if there is a certain new video available in the service.

For each of the statements, the respondents had six different answers to choose from. These were ‘never’, ‘rarely’, ‘quite rarely’, ‘sometimes’, ‘quite often’ and ‘often’.

The second question was ‘Rank the following alternatives after what is most important in a video on demand service.’. Alternatives included:
1. To easy and quickly be able to find the videos I know in advance that I would like to watch.
2. To get inspired and easy and quickly find videos to watch when I do not know in advance what I would like to watch.
3. To easy and quickly be able to continue watching the video where I the last time stopped watching it.

The alternatives had to be ranked with a number between one and three, and they could not share a number. Number one was thought to be the most important while number three would show which was the least important alternative.

The third question was the statement ‘When I know exactly what I would like to watch, then I often find the video by...’. It was followed by a comment box open for any answer the respondent might give.
The fourth question was the statement ‘When I almost (but not exactly) know what I would like to watch, I often know…’. This question was followed by the following nine multiple choice statements:

1. Which genre(s) the video should belong to
2. Which video(s) the video should be similar to
3. Which actor(s) the video should have in it
4. Which feeling the video should give
5. Which director(s) the video should be created by
6. Which rating the video should at least have
7. How the movie poster should look/looks
8. Which year the video should be from
9. Other (followed by a comment box for own alternatives)

The final question was ‘This (if anything) frustrates me when I use a video on demand service (other than internet connection problems and the video content offering in the service).’. A comment box was used to collect the respondents’ answers.

The tool selected for the survey was based on economical factors, the available question types, and the built-in analysis functionality. SurveyMonkey was used since it was a free tool with question types alternatives that other free tools with analysis functionality did not offer.

Following Unger and Chandler’s (2012) recommendations, the distribution channel was selected based on the target group’s typical behaviour. Therefore, the survey was distributed on Facebook, both as a post in the author’s public timeline and also as a post in a group with students. Furthermore, the post was shared by friends and family and therefore, reached more than just the author’s social network. The post contained a link to the survey (see the design of the survey in Appendix D).

The survey was activated for four days, starting on the 31 May 2016 and when it was closed, 60 respondents had answered. The answers were then collected and analysed.

Next chapter will describe the result and analysis of the different research methods used in this thesis. Furthermore, it will present the foundation of the developed guidelines presented in Chapter 6.
5 Result and analysis

Goodman et al. (2012) claim there is no specific correct technique for analysis of qualitative data. The technique selection should be based on which question need to be answered, current limitations and also from the inquisitor’s knowledge and preferences. However, the first recommended step of the analysis is to prepare the raw data for manipulation. It serves as a natural part of the analysis since a natural filtration will occur during the process (Goodman et al., 2012).

The data from the heuristic evaluation and the survey was prepared by being summarized into documents to create a better overview. Moreover, the results from the heuristic evaluation were analysed and several local problems discovered was merged into one global problem. The survey data was collected from SurveyMonkey’s free online tool and transferred to documents to provide a better overview for each questions response.

The first step of analysis of the interviews and the usability tests were to transcribe the recordings. An important part of transcribing the data is to capture what the participants really meant or was thinking about something even if they said something else (Goodman et al., 2012). Therefore, each audio recording was listened to several times to transcribe not only what was said and done but also hesitation, expressed emotions, or other cues of what the person really meant when s/he were answering the question. The same procedure was followed for every video recording from the usability tests to capture movements, face expressions, body language, and other cues that revealed the participants experience. Quantitative and qualitative measurements such as time to complete tasks, completion rates, how the participants performed tasks, errors, recovery, user satisfaction, and emotions were also documented during the transcription. All the interviews and the usability tests were transcribed individually into digital documents.

Goodman et al. (2012) claim there are several advantages and disadvantages with documenting digitally as well as manually and recommends to mix the use of both techniques for the best result. Therefore, all the digitally documented data was printed. When analysed, notes and highlight were added manually. The data from the usability tests and the interviews was even more summarized and sorted into tables to simplify the comparing and analysis. These were also printed and analysed again (see Figure 10).

This chapter presents the results from each of the methods used. Section 5.1 present the usability problems found in the heuristic evaluation. The following section present the results from the interviews while Section 5.3 present the results from the usability tests. Section 5.4 presents the results from the survey. The final section of the chapter presents a comprehensive analysis of all the research.

Figure 10. All data was summarized, sorted, printed, and analysed.
5.1 Heuristic evaluation result and analysis

The heuristic evaluation resulted in a list of 14 usability problems. Three severity grades, from low to high are used to classify the problems discovered. Six of the problems are estimated to be of high severity (see Table 5) and four of them relate to the fifth heuristic of error prevention. Three of the high severity problems are global since they occur in several locations in the system.

Table 5. The six problems of high severity found in the heuristic evaluation.

<table>
<thead>
<tr>
<th>Usability problems of high severity discovered in the heuristic evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Selection of function occurs before user is finished reading the function name.</td>
</tr>
<tr>
<td>2. Hard to see action result or where to continue interaction to reach goal.</td>
</tr>
<tr>
<td>3. No way to undo or go back from faulty or accidental selection.</td>
</tr>
<tr>
<td>4. Time consuming and hard to find a specific episode of a TV show in the display menu.</td>
</tr>
<tr>
<td>5. Non-ergonomic interaction required to reach goal.</td>
</tr>
<tr>
<td>6. Play/pause button display wrong icon (bug).</td>
</tr>
</tbody>
</table>

The first global problem of high severity is that a selection of a function occurs before the user has finished reading and interpreting the function name and icon. Since the describing text of the function is not visible before the users hover with the head gaze cursor over the icon, the user has only 2000 milliseconds, of reading and interpreting the function name before the selection is completed. The problem relates to several of Nielsen’s (1995) usability heuristics, including the visibility of system status, user control and freedom, error prevention, and recognition rather than recall (see Table 3 in Section 3.2 for an explanation of the heuristics).

Another of the global usability problem of high severity is that it is hard to see the selection result, which sometimes include seeing where the user need to continue the interaction to reach his or her goal. One example of this is when the user turns on or off a functionality in the settings menu (see Figure 8 in Section 4.1) but do not see the result of the action since it is outside of the user’s cone of focus that is experienced as very limited to where the button for the functionality is placed since the head gaze needs to be placed over it for 2000 milliseconds in order to complete the selection. The same problem occurs when the user is selecting a button in the main menu but is unable to see the result of the selection that occurs in the display menu at the same time. Since the user is not aware of what change s/he triggered with the selection, s/he also does not know what to do next or how to go back. Therefore, the problem relates to Nielsen’s (1995) usability heuristics named visibility of system status, match between system and the real world, and recognition rather than recall.

The third global usability problem of high severity is that there is no visible way for the user to undo or go back after a selection is completed. If the user has completed a selection by mistake there is no easy way for him or her to undo the selection or go back to the previous state of the system. If a TV show is selected in the display menu, the complete menu will be replaced with episodes from that specific TV show and there is no other way to return to the list of TV shows in the display menu except than for the user to go to back to the main menu and select the TV show button again. The same applies to the Exit VR experience button that will immediately close the application without any control question. This problem relates to Nielsen’s (1995) heuristics considering visibility of system status, user control and freedom, and error prevention.
Another high severity usability problem is that it is time-consuming and hard to find a specific episode of a TV show in the display menu. The episodes are all displayed with the same images and there is no clear division between the different seasons. Under each episode a small text is displayed and the user has to turn his or her head to start the scrolling of the episodes to reach more episodes. Since the text is very small and the scrolling makes the episode images and texts fly in front of the users eyes, it is hard to read which episodes are currently passing by. This causes the user to have to turn his or her head to start the scrolling, stop it by turning the head back to the centre, read which episodes are currently in front of him or her and then continue the scrolling and do the same procedure over again until the right episode it displayed in front of the user. This is both ergonomically uncomfortable and requires much of the user’s cognitive abilities in focusing on guessing how far s/he has to scroll to reach the desirable episode. The problem relates to Nielsen’s (1995) heuristics about user control and freedom, error prevention, recognition rather than recall, and flexibility and efficiency of use.

The fifth usability problem of high severity relates to the fourth problem. Non-ergonomic interaction is required for the user to reach his or her goal. To open the main menu or scroll through episodes in the display menu it requires the user to turn his or her head several degrees further than what is comfortable to do. This problem relates to the Nielsen’s (1995) heuristic of flexibility and efficiency of use.

The final usability problem of high severity is a bug in the system displaying the wrong icon for the play and pause button. Most video application uses the same button for pause and play and depending on what state the video is currently in, the available functionality of the icon is displayed. If the video is currently playing, the button displays a pause icon and the reversed if the video is paused. This could confuse the user and cause him or her to miss the functionality of the button. The problem relates to the Nielsen’s (1995) heuristics concerning consistency and standards, and error prevention.

The problems of medium severity are that information required for decisions is not visible when the user is interacting with the menu. A help function is also missing. Furthermore, two almost identical icons are used for the functionality to rewind 10 seconds and to exit the VR experience, which could confuse the users. Another problem is that the icons look very similar to an icon that it usually used for an undo or back functionality, which can cause even more confusion. Moreover, the user receives different progress feedback when interacting with different functionality in the system. During most interactions, the user will see a circular progress bar showing how much is left before the selection is completed. However, when the user place the head gaze cursor over the fast-forwarding functionality the feedback is only shown in the video and in the function name that changes with the accelerating speed. The final usability problem of medium severity is that icons and explaining texts of the functionality in the settings menu are hard to understand. A complete list of all the usability problems is available in Appendix E.

5.2 Interview result and analysis

The aim of the interviews was to research the participants’ previous experience, behaviours, goals, and values related to VoD services. A second goal was to research their previous experience of, values, and expectations of VR in relation to using it to consume video content. The final part of the interviews was aimed to research if the participants had any previous experience of controlling a system with their eye movements or head movements and if they had any expectations of how that would work.
The interviews revealed that five out of six participants had between three to five years of experience using VoD services. Netflix was the most commonly used VoD service since all participants claim they had or was currently using the service. Five out of six participants mentioned they also had used HBO Nordic and four out of five had used Viaplay.

When asked which technical platform they usually used for VoD services, the answers varied. Half of the participants mentioned their laptop and their smartphone, while two out of six mentioned Smart TVs, Apple TVs, and tablets. Other technologies mentioned by individuals were for example Wii U, PlayStation and a Chromecast connected to a TV. When choosing a technical platform for VoD usage half of the participants answered that it depended on what was easier in the situation and based on their current location. All participants said they were using VoD services in their own home and four out of six claimed they only used VoD services in their home and exclusively in their couch or in their bed. Two persons also said they used it during trips or at work. Four out of six participants claimed they most often used VoD services in company and half of the participants said they were using VoD services as a substitute for linear TV.

When asked how the participants decided what content to watch they often had several factors influencing their choice of video. Four out of six said they usually browsed in the VoD service, checking recommendations, genres, or news. Four out of six also mentioned that it was always based on the situation. Half of the participants mentioned that their mood decided their preferred content at the moment while another participant said that when s/he had time to kill, the choice was always movies and that if s/he was watching in company with someone else it also influenced the choice of video. Half of the participants also mentioned their TV show addiction and that they just kept following their TV shows. Four of the participants also mentioned recommendations or what they had heard about from friends, bloggers, or media.

Important functionality in a VoD service was according to half of the participants to be easy and just intuitively work. Five of six participants also mentioned sorting that is relevant for them. Emphasised was also the possibility to search and find wanted content in a fast way.

Five out of the six respondents claimed that they were easily affected by their surroundings while they were using VoD services and that they were easily distracted. The typical usage scenario was that the participants came home from work and then sat down in the couch and continued watching a new episode of a TV show they were following. Two participants claimed that when they are trying a new movie or TV show they usually judge quite quickly if it is good enough to continue watching. Within about 20 minutes or half of the episode, they will decide if the content is worth to continue watching.

The reason the majority of the participants were using VoD services were because they wanted to relax and disconnect from their lives for a while and at the same time get entertained. They wanted to be immersed in the experience and really feel like they are experiencing the story in the video.

Requirements from the participants to reach that feeling they first of all needed a video with a good story and good video screen quality, sound quality and that everything just works without disruptions. The possibly best experience half of the users could imagine while watching a video was in cinema environment with a big screen, good sound where they are alone and it is a calm and relaxing environment. The most normal thing that frustrated the participants in VoD service was when there were problems with the service but they did not know what the problem was or how to rectify it. They were also frustrated when it was hard to find what they were looking for.

The second part of the interview revealed data about the participants’ previous experience, opinions, and expectations of video in VR. First, the interviews revealed that five out of six participants had
previously tried an HMD with a smartphone but only very shortly. The general experience they had was that it was exciting, cool, and fun.

The participants’ opinions on how they felt about watching a video in VR was collected and five out of six had a positive attitude towards it and thought it could be a good and cool experience. When asked to imagine where and if they would possibly watch a video in VR, five out of six claimed that they would do it at home. Moreover, two participants mentioned watching it while travelling and one person said s/he did not want to do it at all. The situations they would use it in would be similar to normal VoD service usage, including at nights when they want the time to pass by or be entertained. Five of six participants thought that video in VR would have the advantage of adding to the experience by making the participant feel more immersed and easier disconnected from the surroundings. Other advantages that was mentioned was that it would be luxurious to create a personalized space in VR where the video watching would occur and that it would be convenient and nice to watch a video in bed without necessary needing to have a TV in the bedroom.

Imagined negative aspects of watching a video in VR was, for example, the social aspect that would be hard when the person was wearing the HMD, or that the HMD would be ergonomically uncomfortable or that it would be an uncomfortable for the eyes. Two participants also mentioned that the video might not be suitable for children or people with disabilities since the video might get ‘too real’. When asked what they were expecting a UI to look in VR they mentioned that it probably looks like a normal UI but floating in the air and they would probably control it with some kind of handheld device. None of the participants had ever tried controlling a system with eye movements or head movements before and generally they had no idea how it could possibly work.

5.3 Usability test result and analysis

This section present the result and analysis from the task-based scenarios and the following interview in the usability test. When data that is not symmetrical, like task times are being analysed, the mean can be heavily influenced by a few extreme data points and therefore, is not a qualitative measure of the middle value (Sauro & Lewis, 2012). The median then provides a better idea of the most typical value. Therefore, the median value of the task times is used for the analysis.

The first scenario was directly after the participants put on the HMDs and entered the application. They were asked to explore, share their first expressions, and tell what they thought they would be able to accomplish with the application.

The time they spent exploring until they seemed to be satisfied and done varied between two to five minutes but the majority of the participants spent about three minutes before they showed signs of being ‘done’ with the task. Since it was a scenario with a task with no right or wrong or completion rates the most interesting data was first impressions, how much of the functionality the participants found, and how fast and well they understood the interaction method. Also errors, recovery, and general comments during the scenario were valued.

The first impressions when entering the application were almost exclusively smiling faces and the participants expressed feeling amazed, surprised, happy, and overwhelmed. When they started exploring the application several participants showed curiosity, concentration, and they were often surprised. One out of the participants was notably sceptical about everything.

Quantitative data was gathered regarding the amount of the six information areas available in the application (see Section 4.1) and how many was discovered by the participants during the first scenario. No participant discovered all the areas during the scenario but five out of six participants
discovered the display menu and four out of six discovered the play menu. Half of the participants found the main menu and two out of six found the information area. Only one out of the six participants found the settings menu.

The participants were only told in the beginning of the interview before the test how the interaction would work in a simple way that they would control the UI with their head movements but think it was their gaze that would control the UI. Out of the six participants, only one participant was not able to understand how to interact with the UI during the first scenario. The median time to understand the interaction method of head gaze was 55 seconds.

General findings included that two out of the six participants accidentally exited the application without knowing that they did it and the application had to be restarted. A majority, four out of six, participants did not turn their head gaze far enough to the sides and up and down for the information to appear and therefore, missed content. One participant claimed s/he was experiencing simulator sickness when menus appeared in front of him or her and when the scrolling in the display menu started. One person had problems understanding what happened when s/he was interacting with the buttons in the settings menu. Furthermore, two persons expressed that the scrolling feature in the display menu was ‘volatile’ and ‘ponderous’ when the content was massive.

The second scenario was for the participants to imagine that they would like to find the title of the video currently running on the screen. This was considered to be fairly easy by five out of six participants that found the title with the short median time of 6 seconds. However, two of the participants needed a minute to find it and one person was not able to find it at all. The participant, who did not find it, actually found the title information but did not realise that the information was the video currently playing, but thought it was another suggested video. That participant gave up after s/he had been searching for the information for 90 seconds. Later during the other scenarios, the participant realised the information displayed was about the video currently playing on the screen. One participant accidentally exited the experience without being aware of why the application closed down. Two out of six participants were expecting the information to appear on the video screen and first looked there. Moreover, half of the participants also turned to the right before they turned to the left when searching for it.

The third scenario was for the participants to imagine that they would like to start another video. This was easy to the participants and all were able to find and start another video and within the low median time of 9 seconds. The participants went straight down to the display menu and chose a video directly. One of the participants started a video but was not sure if s/he succeeded to reach the goal so s/he looked to the ‘information area’ to confirm that the right video was selected and had started running.

The fourth scenario encouraged the participants to fast forward in the video. This was also observed to be easy for the participants and all were able to perform the task with the median time of 11,5 seconds. Small noticed problems were that two participants had problems opening the play menu on the first try and accidentally opened the display menu instead. The error was very easily and quickly recovered from and no participants mentioned it as a problem. Four out of the six participants needed extra time to realise that the action was successful and was notably concentrated. The automatically accelerating speed of the fast forwarding button was experienced by two users as ‘smart’ and by one as frustrating since the person lost control and wanted to continue with a certain speed for a longer time.

The fifth scenario tested if the participants were able to pause the video that was currently playing. This was also observed as easy and all the participants were able to perform the action with the median time of 5 seconds. The participant that needed the longest time to complete the task was
confused by the bug in the application displaying the play button instead of the pause button even though the video was currently playing. The other participants did not even notice that the wrong icon was displayed.

The sixth scenario was for the participants to start the video again and all participants easily performed the action with the exact same median time of 5 seconds. Many expressed how easy this action was.

The seventh scenario was for the participants to find and start the first episode of the first season of the TV show the Simpsons. This was the scenario where the most participants had problems to reach the goal of the task. At the end, five out of six were able to reach the goal but the median time was two minutes. The participant who was not able to reach the goal gave up after 90 seconds. That participant had then selected the TV show button in the main menu but was struggling to find where to go next and continue the search for the right TV show, season, and episode. That participant wanted to try to perform the task later during the last scenario of the test when s/he had discovered more about the application but then gave up after three minutes when trying to find the right episode scrolling through the display menu with all the episodes. The participant explained s/he was experiencing simulator sickness due to all the scrolling and flying objects.

Five out of six participants started their search by scrolling through the current content in the display menu. One participant selected the TV show button from the main menu and then turned his or her view down to the display menu and before s/he realised it, the participant had selected a TV show by mistake. This causes the application to replace all the currently displayed TV shows in the display menu and instead display the episodes of the selected TV show. The participant was confused and tried to find an undo or back functionality to go back to the previous view. In the search s/he discovered the icon used to exit the VR experience in the main menu and moved the head gaze over it, with the belief it would be an undo button. However, the participant realised last second the action the button would cause and removed the head gaze from the button before it was selected. S/he continued to search for an undo button in the play menu and there found the orange circle in the middle and tried to select it. Finally, the user got frustrated and said ‘How do I get out of here?! Okey, I am stuck! I guess I have to start all over by selecting TV shows again.’

Three out of six participants had problems seeing how the selections in the ‘main menu’ controlled the content in the display menu. One participant found the button for TV shows in the main menu but commented ‘Oh, this list will be huge so I will try to find search instead!’

Five out of six participants had problems when they were scrolling through the display menu. They had to pause the scrolling and go in to the centre of the display menu to be able to read the episodes names and stated that they were having trouble to see what they were scrolling past. Half of the participants scrolled past the episode they were looking for without even noticing it. Comments of illustrating the users experience included ‘Oh my God – you have to turn so much to scroll!’ and ‘There must be a smarter way to do this?’ One participant also had trouble realising how to stop the scrolling once it had started. Another participant tried to speed up the scrolling by interacting with the display menu progress bar underneath the display menu.

The eight scenario asked the participant to find the TV show Mad Men instead. This was observed to be easy and all participants were able to complete the task with the median time of nine seconds. The participants claimed it was a simple task and easy because they just did as they did before. One person needed a longer time to complete the task and this was since the person spent that time scrolling through the display menu to find the right TV show.
The ninth and final scenario the participants were presented to was to continue exploring the application and find as much functionality as possible and tell what they thought all the menu options and UI elements would lead to. This scenario revealed problems user had after a few minutes of using the application and how well they found and understood everything in the UI.

When two of six participants stated that they had found everything they thought there was to find in the UI, they still had not found the settings menu. The moderator then told them there was still an unexplored part of the UI. As a result, they were then able to find that menu as well.

Five out of six participants had understood the main menu and its functionality except the exit button. It was one of the biggest issues in the UI and four out of six participants had not memorized the functionality of the exit button even though they already had used it hovered over it before. One person claimed it was the ‘back button’ even though the person had selected the button earlier during the scenarios. Another person claimed that the rewind 10 seconds button in the play menu, with almost the exact same icon as the exit icon, was a back button. Five out of six participants did not understand the channel button in the main menu until they read the button name while hovering with the gaze cursor over it. Moreover, two out of six participants had not yet memorized where the video title and information was displayed and was reminded when they were searching through the UI and turned their heads to the right. Another interesting observation was that all the participants tried to interact with UI elements that had no functionality, such as the orange buttons in the centre of the settings menu and the play menu along with the display menu progress bar. Only one participant understood what the display menu progress bars purpose was.

Another interesting discovery was that all the participants had trouble understanding the functionality of the buttons in the settings menu. Four out of six participants kept turning functions on and off several times and showed high concentration when they tried to discover which change the button selection had caused. Half of the participants finally, understood the functionality of screen type and the button scene. Moreover, two participants understood the function of the leak light. However, one participant thought the button changed the colour settings of the video. Only one participant was able to understand the SFX button that turned on and off button sound effect in the application and two participants were very confused since they thought they would have three bars as displayed in the icon where they could control the sound of the video.

One person had not figured out how to enter the view mode and was asking about how to turn off all the functionality and have a larger screen. Another participant was annoyed by how every little movement of the head made the whole UI move and said s/he wanted to turn that off while watching a movie. Another participant claimed s/he had trouble seeing which functionality was activated and which was not. Moreover, another commented on the interaction method and said ‘You have to go through quite a lot of trouble to interact with your head like this just to start a video.’

Another observation was that two out of six participants kept holding the HMD with their hands and changing the hand to the other when they got tired in one arm. They claimed the HMD was heavy and pressing on their nose. All the participants were noted to have a change in voice from a normal tone to a more nasal tone confirming the pressure the HMDs caused on their noses.

After the test, the participants were asked several questions about their experience. Generally, they said that the experience was cool, impressive, fun and that they were feeling immersed. But they also said that they were feeling stress and confusion. Three out of six participants claimed that they think watching a video with an HMD in VR would give a better experience than a normal TV screen. One of the participants who thought a normal screen provide a better experience said it was because of the limited resolution and that the distance from the video screen was not optimal in the prototype. The other participants said that s/he saw no extra feature in the prototype that s/he would not be able to experience with a normal TV screen. For the last participant, it depended on the
video content. S/he claimed that action or nature videos would be better with an HMD in VR but other content would be better experienced on a normal TV screen.

The things that the participants wished to experience in VR included movies, places, and games. One participant was very positive and claimed ‘This! Exactly this – just movies and nothing else’.

Moreover, the navigation of the application was experienced both positively and negatively. The negative experiences participants shared was that it felt a bit ‘annoying’ to use for longer periods of time. One person also mentioned that s/he thought it felt like s/he looked very dumb while wearing the HMD and interacting with the VRGUI. Two participants claimed they would have preferred to use the interaction method in combination with a handheld control while one participant claimed s/he felt no need of any other control and was positively surprised by the fact. One participant said it was easy and intuitive to use and no functionality was missing while another participant claimed s/he was missing a search function. Moreover, one participant liked the idea behind the interaction and the animation of the menus when they appeared.

When asked if the participants felt like they could control everything they wanted to control, three out of six claimed they could but said that the settings menu was experienced as complicated. Two of the participants who answered no said that they had trouble with the settings menu and that they were missing feedback if they had performed a selection or not. Another participant told that s/he had to try many times to succeed with a selection and that there were many accidental selections that occurred.

The dwell time that was required for the participants to perform the selections was in general experienced as a good amount of time. One participant said that the time felt too fast in the beginning but then felt too slow as soon as the person had realised how the interaction technique worked. Also, another one claimed that s/he would probably have wanted the dwell time to be shorter when s/he got more experienced using the interaction method. One participant mentioned that the dwell time for selecting a movie was perfect but for the buttons with unknown functionality it was too fast and it caused a stressful situation. One person claimed it was a problem when s/he was just looking at the different movies and accidentally performed a selection of a movie.

When asked if something was hard to understand, see, or perform four out of six participants mentioned the settings menu and the functionality in it. Moreover, two participants mentioned it was hard and time-consuming to find a specific episode of a TV show. One participant described that s/he had not found a button for volume control and one stated that it was hard to see where s/he was at the moment and that s/he was forced to go back and redo things many times. Another participant mentioned the resolution and that it was hard to read the episode names while s/he was scrolling. Finally, another participant described it to be frustrating trying to interact with the orange button in the centre that was not possible to interact with.

When asked if something was easy to understand, see, or perform half of the participants mentioned playing and pausing a video. One participant described that it was easy to find information about the video that was currently playing and another mentioned performing selections since the circular progress bar was very easy to understand and follow. Another participant mentioned that it was easy to see the videos in the display menu and another person mentioned that the main menu was easy to use.

Two out of six participants were surprised by the quality of the virtual world. Other things that surprised the participants were when they found new functionality or when they by accident exited the experience.
Half of the participants said they desired and was missing a search functionality in the VRGUI. However, two participants said they were not missing any functionality at all. Two persons mentioned that it would have been desirable with another way of listing the episodes with a clear separation between the seasons of a TV show. One person claimed s/he felt limited in the way s/he could interact with the system. S/he said ‘Sometimes I feel limited using a touch screen because I can not right click as I usually do when I am using a computer mouse or trackpad. Here I have even more limits and it makes me feel a little handicapped. It felt like I had lost my arms!’

Things that frustrated the participants were for example that the HMD was heavy, it took a long time to reach certain goals in the VRGUI and that there was no search functionality. Half of the participants mentioned that it was a lot of black space around them in the virtual world and that they would have preferred to have the video screen covering the full visual space. One person mentioned that it was frustrating to not know what all the functionality was and how it worked. Another person mentioned it as disturbing that the HMD was so sensitive to every little intended or unintended movement the participant did with his or her head.

When asked what could have improved the experience, half of the participants once again mentioned having the video screen covering the whole virtual environment. Another participant mentioned better adoption of the distance of where the display was shown to the participant. Moreover, one person wished for a clearer VRGUI so s/he could more easily understand all the functionality and a clearer fast-forwarding function. One participant also wished s/he did not have to turn and read texts to find a specific episode. Furthermore, another participant described a desire to have more of a finishing touch on the VRGUI elements to make it look a bit more ‘cool’ and more ‘playful’. That person would have liked the VRGUI to look a bit more hand made, lavish, and polished.

When asked if there was anything that would have made the participants choose to watch a video in VR instead of on a regular screen they claimed that they would choose it since it is convenient, comfortable and also adds to the experience. One participant said s/he would always choose the VR experience for video watching and two persons said nothing could make them choose watching a video in VR instead of watching video on a regular screen. One person mentioned that it would be convenient to be able to watch a video undisturbed in a room full of other people and that it was really comfortable to be able to do that and to still see a big screen and have a great experience.

Things that the participants generally liked about the VRGUI were the display menu that was experienced as intuitively placed and stylish. Furthermore, participants liked the interaction method including the dwell time used for selections. One participant even said ‘that you can control the UI with only the thought! Or it is the head or...how do you explain it?’. Another said that ‘The dwell time was great! You could easily see when the progress circle was loading and turn away if you wanted to. That was really great – I think that was the best part!’ Moreover, one participant liked the colours and the icons and another mentioned the video screen that was looking ‘clean and stylish’. Another mentioned the crosshair used for the cursor and said that it was discreet enough to not interfere with the experience but at the same time visible enough to be noticed and used.

Things that the participants disliked were, for example, the settings menu and the display menu. The settings menu was experienced as hard to understand and the display menu was experienced as volatile. One participant mentioned how it was hard to get an overview of the movies and the TV shows and that it was hard to orient in the VRGUI. Another participant described how s/he did not like that some UI elements were not possible to interact with. Moreover, one other participant suggested that the dwell time should have been adapted to different lengths for different selections. One participant stated that the interaction technique was not sustainable and would not work in the long run. S/he meant that the interaction method would work in the beginning or a few times but when used on daily bases it would be too demanding to use.
The participants were also asked what they would have changed to make the VRGUI easier to use. Two of the participants described using another interaction method including some kind of handheld or hand-controlled device. This while two participants mentioned adding search functionality. One participant desired something that easier linked the surfaces the participant had to find when performing an action. Moreover, one participant also mentioned adding a back functionality. Finally, a participant said s/he wanted a larger play button in the middle and the possibility to turn off the rest of the VRGUI while watching a video.

When asked how the participants would have changed the VRGUI to make it more fun to use, half of the participants mentioned that the virtual world was very dark and maybe a lighter environment or the possibility to change to other environments would have made it more fun to use. Furthermore, one participant said adding game elements to the interaction and creating more of an artistic experience while interacting. One participant mentioned displaying videos in a grid of five times five where different rows could be used to sort different videos or maybe display the videos in another type of circle away from the user instead of around the user in the current version.

5.4 Survey result and analysis

With the aim to research the users’ behaviour, preferences, and frustrations related to VoD service usage the survey revealed UX data of relevance for the specific application.

The first question, with sub-questions, was aimed to research the frequency of different goals the user can have when they are starting a VoD service. Different scenarios of goals were presented, and the responders’ task was to answer how frequently those goals were the reason they started the service. The alternatives they had to choose between were ‘never’, ‘rarely’, ‘quite rarely’, ‘sometimes’, ‘quite often’ and ‘often’. Chart 1 displays the result from the first question in the survey that is explained in the following paragraph.

![Chart 1. Result from the first question of the survey showing the amount of responses (n=60) for each frequency alternative for each different goal.](image)
The first goal the respondents was asked to answer the frequency about was how often they open a VoD service with the goal to find something to watch when the user does not know what to watch in advance. Out of the 60 responders, 18 persons said they did it sometimes and 18 persons said they did it quite often. When asked how often the respondents open a VoD service with the goal to find something to watch when they almost know what they would like to watch in advance, 24 persons answered quite often and 18 persons answered sometimes. Another goal was how often the respondents opened a VoD service when they knew exactly what they would like to watch in advance. Out of the 60 responders, 18 persons answered often and 15 persons answered quite often. When asked how often they opened a VoD service to continue watching something they started watched earlier, 23 persons answered often and 13 persons answered sometimes. The final goal the respondents were asked about was how often they open a VoD service with the goal to control if a certain new video is published in the service. Only 4 persons answered often and 19 persons answered rarely while 11 persons answered sometimes.

The second question in the survey asked the respondents to rank out of three alternatives, which was the most important in a VoD service. The three alternatives were ‘To easily and quickly be able to find the videos I know in advance that I would like to watch.’, ‘To get inspired and easily and quickly find videos when I do not know in advance what I would like to watch.’ and ‘To easily and quickly be able to continue watching where I stopped watching in a video during my last use of the VoD service.’. This question only had minor differences in numbers and no clear pattern. Out of the 60 respondents, 22 persons ranked the alternative to continue watching as number one and 22 persons ranked to easily find a video when they know in advance what they would like to watch as number two. A total of 27 persons ranked the alternative of getting inspired as number three in the priority. However, it turned out to be a misunderstanding of how the ranking scale worked and therefore, the numbers might be misleading. Some might have ranked their most important as number three, and others might have ranked the most important as number one. Therefore, this result is excluded from the analysis and not visualised.

The third question was designed with a comment box for open answers. The question was ‘When I know exactly what I would like to watch, then I usually find the video by...’. A total of 59 responses were collected since one respondent decided to not answer the question. The responses were summarized in a document and analysed to find patterns of similarities in answers to discover patterns. Some respondents answered several ways they usually found the video. A majority of the respondent, 44 persons, answered they would use a search functionality. Within that group of respondents, some wrote that they would search for the title, other wrote that they would search for the genre, and others wrote keywords. Among the rest of the responses eight was related to just finding it directly on the start page of the VoD service and five persons answered they would look for it by genre. Chart 2 visualise the result of question two from the survey.
The fourth question was a question with multiple choices where the respondents were asked to answer what knowledge they usually have when they almost, but not exactly, know what they would like to watch. Respondents could select several options from a list or add their own under ‘other’ if they usually had information that was not in the suggested list of alternatives. The complete list of alternatives is previously described in Section 4.4. As Chart 3 shows, out of 60 responding persons, 54 persons often knew what or which genres the video should be. Other popular responses were that 24 of the responders usually knew which feeling they wanted the video to give them and 23 persons knew which video or videos the video should be similar to. Furthermore, 22 of the respondents also usually knew which actors should be in the video.

The final question of the survey was designed with a comment box to collect qualitative data and opinions. The question was ‘This (if anything) frustrates me when I use VoD services (except internet connection problems and the video offering in the service). Out of the 60 participants,
only 33 responses were collected. Several did not answer or answered ‘nothing’ and some answered things relating to the exceptions of the question. These answers were excluded from the analysis. Out of the 34 responses, seven persons answered that not being able to find something to watch or having to scroll through a lot of content was the most frustrating thing (see Chart 4). This while six persons answered issues regarding the subtitles and the settings of them. Six persons also mentioned bad recommendations in the service. Four persons answered it being hard to find the way back to something they already started watching. Other less frequent responses included a bad overview of content, the auto stop functionality, that it was hard to reach information about a video, country restrictions in content and different types of UI keyboards that were hard to use.

5.5 Comprehensive analysis

The aim of this section is to describe the comprehensive final analysis procedure and how the different data was compared and analysed together. The goal of the step was to find patterns in the data and create groups based on the found patterns. A bottom-up approach was used where the data was further explored to discover natural groups (Goodman et al., 2012). First, each method’s result data was analysed individually and then compared and analysed comprehensively.

5.5.1 Identifying the users’ goals, needs, behaviours, preferences, and context of use

The aim of the research was to create a better understanding of how VRGUIs controlled by head gaze should be designed for qualitative UX. Furthermore; the research was limited to the application of VoD services. Since qualitative UX is defined as the fulfilling of users’ pragmatic and hedonic goals, one important part of the analysis was to identify those goals and other subjective aspects influencing the users’ experience. Therefore, the first part of the analysis focused on identifying users’ goals, expectations, needs, behaviours, preferences, and context of use.

The interviews revealed that participants are most often using VoD services with the motivation and higher goal of relaxing, disconnecting from the real world, and to get entertained. These are interpreted as the hedonic goals of the users. The pragmatic goals of highest priority included being able to continue watching a video where the users had paused it during the previous use of the service and to easily find specific videos the users knew beforehand that they wanted to watch. Secondary pragmatic goals were to find a video when the users almost knew what they would like to watch and to find videos when they did not know beforehand what they wanted to watch. The three hedonic goals and the three pragmatic goals of highest priority is visualised in Figure 11.

![Figure 11. The users’ pragmatic and hedonic goals when using a video on demand service in virtual reality.](image-url)
Expectations of video in VR include an even more immersive and ‘cooler’ experience and that the user easily forgets about the surroundings. Furthermore, many are concerned if the experience will be comfortable, both for the eyes and also ergonomically on the head. Moreover, there is a general concern of ‘looking dumb’ while wearing the HMD. This may be connected to the image the users have of standing with the HMD and waving their arms. In the virtual world, many are not sure exactly what to expect which makes them curious since it is something new and most can only visualise a VRGUI similar to a normal GUI but floating in the air in front of them. Most also imagine the control to be made with a handheld device, probably based on previous experience of how UIs usually are controlled and maybe also since the HMDs have mostly been marketed for gaming purposes.

The physical location was analysed to most likely be in the users own home, in their couch or their bed. Also traveling by car, train, and airplane could be potential locations for usage. The situation is usually at night, after work when the users want to relax and think about something else or when they are travelling and want time to pass by.

Many users are impatience to find what they are looking for and get annoyed if they have to scroll too much. Since the context of use is usually after work at night, many users are expected to be tired and only willing to use minimal effort to reach their goal. Many users are also easily distracted and want to get entertained and immersed as quickly as possible.

The best video watching experience is when the user is alone in a calm and relaxing environment with a big screen with good quality and good sound. Preferences for finding information include using search functionality and easy find videos based on genres, feeling the video lead to, similar videos and actors in the video. The most frustrating scenario to the users is when it is hard to find the video they are looking for.

5.5.2 Analysis of the discovered user experience problems

The second part of the analysis was to analyse the UX problems found in the data from the heuristic evaluation, interviews, usability tests, and surveys. The bottom-up approach was used to discover patterns. Global UX problems counteracting with the pragmatic and hedonic goals of the users were the problems where most focus was placed.

The majority of the UX problems discovered seemed to relate to the fact that users were lacking information and control in the VRGUI (see Figure 12). Several reasons were found to why they were the big UX problems of the application and they are further described in the following paragraphs. Even though they are listed as two separate parts, they are strongly connected to the users.

![Figure 12. User experience problem discovered pattern.](image-url)
The first big UX problem of the application is that users have a shortage of information. Patterns discovered in the data reveal that this is because the users are not presented with enough information, the information is located too far away, it is too hard to find and too hard to understand.

The information in the prototype has a limited visibility. Only one of the information areas in the prototype is available at the time. For example, when the user is exploring a new video in the display menu, s/he cannot see the information area to the right of the video screen with information about the video that is currently running on the video screen. This loads the cognitive abilities of the users since they have to keep information in their mind. Another example of this is that the explaining text to each button of functionality is only visible when the user holds the head gaze over that specific button. Furthermore, when the text is displayed, the user has less than 2000 milliseconds to interpret the text and decide if the button is the desired action before the selection occurs. If the action is not, the user then has to quickly turn his or her head gaze away from the button.

An example of when the information is too far away is when the user is trying to look for a specific video and therefore, turn to the main menu, select for example TV shows and then can not see where to continue looking for the specific video. This since the display menu is not directly in connection with the main menu and the users cone of focus is observed to be very limited and small during the interaction with a UI element. The user has to focus and concentrate on the selection itself since the user is required to retain with the head gaze cursor over the button for 2000 milliseconds for the selection to occur. Therefore, the user’s cone of focus is noted to be extremely limited to the immediate area around the UI element in focus. Another example of when the information is located too far away is with the menus themselves. Several users were observed to turn their heads numerous times when exploring, but not far enough in order for the menus to appear. Two users even missed the setting menu until they were told about it at the very end of the usability test.

Another example is when users are fast-forwarding in a video and the action requires the users to keep the head gaze over the button, which by the same time leaves the timeline of the video out of the users cone of focus and FOV. The user then has a shortage of information that might be required for the user to make decisions during the interaction, forcing him or her to stop the action, control the information and then go back and continue the interaction to correct the result.

These examples also prove that the information is too hard to find. An example of this is when the user is trying to find a specific episode of a TV show in the display menu. Since some TV shows have many seasons and episodes, and they are all displayed in the same menu with no difference between them except the very small title text below them and the user is required to scroll through them to find the episode it turns into a very time-consuming task. The users cognitive abilities are highly loaded since the user will try to read the small texts while scrolling in order to stop the scrolling at the right moment and select the desired episode.

Moreover, the information is too hard to understand since the use of icons and describing text does not match with the users language concerning Nielsen’s (1995) usability heuristics of match between system and the real world and consistency and standards. The result of the information being located too far away and being hard to find is that the users are having an uncomfortable experience having to turn their heads repeatedly and further than the comfortable 30 horizontal degrees according to Chu’s (2014) recommendations previously described in Section 2.3.

Another part of the information problem is that information is hard to understand. Several users did not understand numerous icons and texts explaining the functionality in the application.
Furthermore, the functionality was not clear even when the functionality was activated since the information was located too far away from the user’s cone of focus.

The second major UX problem, tightly connected to the information problem was that users did not have complete control over the VRGUI. This mostly related to the interaction method, since it is designed so selections occur automatically, even if the user is not aware of it and only places the head gaze over a button by coincidence. This is an example of the Midas touch problem, previously described in Section 2.4. Since several accidental selections occurred during the tests, many users find themselves disoriented and not completely aware of the system’s current state. Moreover, this is also tightly connected to the information problem and how the user has to move the head gaze around to the different information areas, one by one, to figure out where s/he is located and what functionality is activated.

As previously mentioned, the dwell time before a selection occurs is predetermined to 2000 milliseconds and therefore, the user do not have much time to read, explore, and process the information in front of him or her. This causes a very stressful situation where the user is afraid to explore the content since s/he does not want to cause any accidental changes in the system. Connected to this is also the issue that the user has no visible way of undoing or going back from an accidental or unwanted selection. Since it is not completely clear to the user what s/he can interact with and what will cause changes to occur in the system this also adds to the user feeling confused and stressed. Moreover, another example is that when the user selects the button for exiting the experience, there is no control question or any additional step before the whole screen turns black and the user has to remove the smartphone from the HMD and start the application again. All this leaves the user feeling like s/he does not have full control over the system.

Finally, the UX problems found were analysed in relation to the users’ goals, expectations, needs, behaviours, moods, preferences, and context of use. In order for the user to accomplish their hedonic goals of feeling relaxed, immersed and entertained, the experience of the VRGUI must be simple, fun and stress-free when they are trying to reach their pragmatic goals. One part of this is creating a VRGUI with a system architecture that focuses on making the most important pragmatic goals of the user easy and quick to reach. This could include keeping the previously watched video directly on the ‘start page’ of the application in order for the users to easily start off where they previously stopped watching a video the last time. It could also include keeping search functionality within easy reach and a possibility for users to search for videos based on genres, feeling the video will lead to, similar videos, and actors. Moreover, guidelines developed from the result and analysis of the research is described in the following chapter.
6 Suggested guidelines

The research question of this thesis project was what aspects should be considered when designing a graphical user interface (GUI) in virtual reality (VR) - controlled by head orientation input - for a qualitative user experience (UX). Since the application the research question has been limited to is video on demand (VoD) services, the research has also focused on gathering UX influencing data about VoD services.

Based on the results and analysis of the research, previously described in Chapter 5, and with respect to the background literature review in Section 2.7, seven suggested guidelines have been developed and are presented in the following paragraphs.

Generally, it is recommended to always design GUIs following usability guidelines so users in an easy and satisfying way can reach their pragmatic goals. At the same time, the design of the GUI should also allow the user to fulfil the users’ hedonic goals as soon as possible. This includes keeping the pragmatic goals of the users achievable within as few selections as possible and also easy to find. For a VR application in VR according to the research, this could be to keep previously watched videos on the start view of the application and a search functionality, following the users’ needs and preferences, where they can easily find the video they are looking for. However, these recommendations are specific for the application of VoD services and are not the main focus of this thesis. The following guidelines are developed for the application of VoD services but are believed to be more generally applicable for other applications and contexts and therefore, serve towards answering the research question in a broader sense.

1. Place the graphical user interface so it is comfortable to explore and interact with

In VR the GUI has unlimited possible placements in three dimensions. When compared to designing for digital 2D screens where the user usually is able to see the whole screen at once and easily get an overview, the user can not usually see the entire virtual environment at the same time.

The maximum field of view (FOV) of the human visual system, as previously explained in Section 2.3, is approximately 200 degrees (Bowman, 2004). However in head-mounted displays (HMDs), the FOV is about 94 degrees (Oculus VR, 2015c). Users can change the orientation of their heads to explore more of the virtual world. According to Chu (2014), users’ can turn their heads comfortably 30 degrees to the left and right side, 12 degrees down and 20 degrees up. Moreover, the depth perception is also important in virtual environments. Objects should be placed at a distance between 75 centimetres (Oculus VR, 2015b) and 20 meters away from the user (Chu, 2014). Following an older recommendation for the minimum distance, Mike Alger (2015a) suggests a main content zone that is displayed within 155 degrees horizontally and at a distance between 0,5 meters to 20 meters away from the user (see previous Figure 6 in Section 2.3). However, there is a difference between content and a GUI, and therefore, these distance suggestions are replaced.

The recommendation of placement to make the UI comfortable to explore follows the degrees suggested by Alger (2015a), the minimum distance suggested by Oculus VR (2015b) and the maximum distance suggested by Oculus VR (2015a). This leads to a recommended placement within 155 degrees horizontally, at least 75 centimetres away from the user but not further than three metres away in the virtual environment.

More importantly is that the elements of the GUI that the user should interact with or needs to place the gaze cursor over should be kept within the comfortable interaction zone. The suggested
comfortable interaction zone is based on the ideas of Chu (2014) about how people can comfortably turn their heads. The centre of the zone, here referred to as the centre point, is where the user’s head gaze is naturally positioned when the user looks straight forward with the head in the most comfortable natural orientation. Surrounding the centre point, the zone is based on 60 degrees horizontally, evenly distributed to the left and right side from the centre point and 32 degrees vertically, where 12 of the degrees are down from the centre point and 20 degrees are up. Furthermore, the comfortable interaction zone follows the same distance recommendations for the comfortable exploring placement (see Figure 13 for a visualisation of the suggested comfortable interaction zone).

![Figure 13. The suggested comfortable interaction zone visualised. The zone illustrates where user interface elements that can be selected or interacted with in a virtual reality graphical user interface controlled by head orientation input and head gaze should be placed.](image)

While the recommended comfortable view and comfortable interaction zone is based on literature from the background study of the thesis, the guideline is a result of the research in this thesis. Although the guideline’s name focuses on comfortable exploration and interaction, the reasons to follow it exceeds solely ergonomic and visual comfort. Several of the user experience, UX, problems discovered during the research that lead to users’ shortage of information and control (previously described and presented in Section 5.5 and Figure 10), relate to Nielsen’s (1995) usability heuristics about visibility of system status, user control and freedom, recognition rather than recall, and flexibility and efficiency of use. Examples included that users did not find all parts of the VRGUI since they tended to turn their heads far enough in order for the VRGUI to open up and become visible to them. Since the whole VRGUI was not visible at the same time, the cognitive load was heavy and this while the users had to change the orientation of their heads in non-ergonomic ways to interact with the VRGUI.

If the guideline is followed, it helps users to easily find, explore, and interact with the VRGUI in an ergonomically, visually, and cognitively comfortable way while supporting the user to reach both the pragmatic and hedonic goals. Furthermore, it will help users to use the HMDs in contexts were they want to limit the movements required. For example, when travelling, they might sit next to another person, who should not be bothered when they are interacting with the VRGUI and also the user should not hesitate to use the HMD of fear of looking ‘dumb’ while turning his or her head.
This guideline is closely related to three of the guidelines presented in Section 2.7. The guidelines it relates to is ‘Allow users to be lazy’, ‘Use “magic” interaction’, and ‘Place VRGUI at a reasonable and comfortable distance’. Table 1 in Section 2.7 shows an overview displaying which literature describe and support the guidelines. However, this guideline provides a new, and to the author’s knowledge unique, recommendation of a comfortable interaction zone for the head orientation input and head gaze interaction method.

2. Place visual feedback to selections within the immediate interaction area

The interaction method that uses head orientation input to project an approximated gaze cursor in the centre of the user’s FOV, has in this thesis been referred to as head gaze. Even though head gaze is controlled by how the user turns his or her head, the user experiences it as if the cursor is controlled by the direction of the eye gaze. One illustrating example is that the user could turn his or her head in a certain direction and then explore the area in the virtual environment with the head orientation fixated while relocating the eye gaze only. This would leave the cursor fixated while the user explores the surroundings but since the user experiences that s/he is controlling the cursor with his or her eye gaze direction, this might not be considered as an option by the user. Therefore, the user tends to focus with the eye gaze at the same location where the head gaze is currently located to make sure that the selection of the desired functionality is completed. The eye gaze is also necessary to keep nearby the head gaze cursor for the user to make sure the head gaze controlled cursor stays located over the UI element the user want to interact with. Since the user is focusing and concentrating on the object s/he wants to interact with, the surrounding area in the VRGUI becomes out of focus and turns nearly invisible to the user during the interaction.

Therefore, it is necessary to place visual feedback to a selection within the immediate interaction area where the user is currently focusing. The approximated area where the feedback should be placed is within the double diameter of the cursor being used (see Figure 14).

This guideline is developed based on research result and analysis in this thesis. During the research, several participants did not understand the functionality of on and off buttons in the VRGUI since they were not able to see the change that occurred when they interacted with the buttons. Furthermore, several were not able to reach their pragmatic goals since they did not know where to continue the interaction with the GUI after they selected a certain button in a menu. This lead to a shortage of information and control previously described and presented in Section 5.5 and illustrated in Figure 10. Furthermore, it relates to Nielsen’s (1995) usability heuristics about visibility of system status, match between the system and the real world, user control and freedom, error prevention, recognition rather than recall, flexibility and efficiency of use.

Figure 14. The approximated area, double the size of the head gaze cursor, where the visual feedback from a selection should be placed in order for the user to see it. The first image from the left show the circular progress bar of the head gaze cursor during a selection and the second image show the estimation of how the small cone of focus make the surroundings blurry while the third image has marked the area where the visual feedback should be placed in order for the user to see and discover it.
If the guideline is followed, it helps the user to easily see the result of their actions and where to continue the interaction to reach their pragmatic goals.

3. Keep information dense areas interaction-free

With the head gaze interaction method, the interaction with and selection of a UI element starts as soon as the user places the gaze cursor over the element. After the gaze cursor is placed, the user has a limited time to decide if the selection is a wanted action or not since a predetermined gaze time will trigger the selection of the element. In the research, this caused several UX problems such as the users making accidental selections since they were not finished with interpreting and understanding the functionality a selection of the UI element would lead to. Furthermore, this caused a stressful situation where users turned their heads from element to element since they were afraid of staying with the head gaze for too long and performing an unwanted selection. This limited their curiosity, made them feel stressed and several times they were not finished with the exploring of the objects and element before a selection had occurred. This is an example of the Midas touch problem, previously described in Section 2.4.

This UX problem also relates to the user not realising they can move their gaze from one spot in the VRGUI. Therefore, the developed guideline suggests that information dense areas should be kept interaction-free. If the parts of the VRGUI that contain extensive information and content is kept interaction-free, the user will have the opportunity to explore the information in a relaxing way, one of the hedonic goals with the use, and then be able to interact and select a certain content when s/he is ready.

The suggested guideline is therefore to maintain information dense areas interaction-free. If this is not possible, an interaction-free area should be placed in the immediate area surrounding the element. Specifically for information dense areas or areas that are harder for the user to understand, they need more time to interpret and explore the information. To minimize the cognitive load of the user trying to interpret and read information while keeping track of a moving progress bar for the selection at the same time, the area should be interaction-free. This relates to Nielsen’s (1995) usability heuristics about user control and freedom, error prevention, and flexibility and efficiency of use.

4. Use dwell times of various lengths

With head gaze interaction, a dwell time is used to trigger selections in the VRGUI. In the prototype used for part of the research, the dwell time used for selections was constant across different parts of the VRGUI. Related to guideline 3, the user sometimes need more time to interpret or explore the information before s/he interacts with it and the solution in guideline 3 was to keep information dense areas interaction-free but an additional recommendation could be to consider using dwell times of varying lengths.

By allowing the dwell time for different selections in the VRGUI to differ, it could work towards improving the usability, related to Nielsen’s (1995) usability heuristics of user control and freedom, error prevention, and flexibility and efficiency of use. Parts of the UI that require more time to be read or interpreted by the user should have a longer dwell time for selections. Furthermore, the dwell time could be adjusted depending on the users knowledge and experience of the VRGUI. An experienced user might want shorter dwell times while a novice user needs longer dwell times.
5. Avoid using time-limited information

In the tested prototype, the VRGUI was kept minimalistic and near to text free. If the user did not understand the icon of a button s/he was required to place the gaze cursor over the button in order to see the explanatory text for it. However, the hovering information was only available for the limited time before the selection was triggered. If the user was unable to understand the icon from the beginning, and then moved the head gaze over it and was not able to finish reading the hovering explanatory text while keeping track of the selection progress bar, s/he ended up selecting the button. Some users tried to prevent the unwanted selection by moving their head gaze quickly over and away from the button to extend the reading time they were offered. This caused the users to feel stressed, confused when they performed a selection and was not even sure what they had selected. Furthermore, the quick rotations for moving the head gaze was observed as non-ergonomic and the users were notably concentrated since they were required to try to read the text while keeping track of the selection progress bar.

The developed guideline is therefore to avoid time limited hovering information. If the users are presented with the information without having a time limitation to interpreting it, this is believed to reduce the stress, the cognitive load and furthermore, the non-ergonomic head rotations.

The UX problems are related to Nielsen’s (1995) usability heuristics of user control and freedom, error prevention, and flexibility and efficiency of use. If the guideline is followed, it will reduce the cognitive load of the user and allow a stress-free interaction where undesired selections are prevented.

6. Never force users to interpret information in movement

When users are forced to interpret information that is currently moving in front of their eyes, it leads to a high cognitive load where they need to be highly concentrated and focused on the task. Furthermore, this can cause simulator sickness. In the research, only one out of the six participants experienced simulator sickness. Even though it was not ranked as a common problem, it is still extremely negative for the UX and therefore, considered important to reduce.

This relates to the previously described guideline to ‘Avoid rapid movements, do not drop frames or lose head tracking’ in Section 2.7. However, this developed guideline extends the guideline to avoid forcing the user to interpret information in movement. The UX problems closely relate to Nielsen’s (1995) usability heuristic of the visibility of system status.

7. Use standards and affordances to minimize the cognitive load

Minimalistic GUI design is a trend limiting the amount of information displayed to the user. Visual clutter should be avoided to minimize the cognitive load but when too much information is removed, hidden or icons are too simplified, the cognitive load of the user will increase. In GUIs for screens, this is often easily solved by the user hovering to see an explanatory text or by the user clicking on the element to explore what the functionality of it means. However, since an undo or back action is not as easily performed in a VRGUI controlled by head gaze and since the hovering over a UI element that the user can interact with is time limited to the dwell time for the selection, the VRGUI is not as easy and quick for the user to explore if they do not understand the functionality of a UI element from the beginning. Therefore, the suggestion is to use explanatory information of the UI element functionality that is as clear as possible to the user, requiring a minimal cognitive effort for the user to interpret. This usually includes clear icons and short but descriptive titles for the functionality. Moreover, this relates to the previously described guideline to ‘Help the user by using affordances’ in Section 2.7.
7 Conclusion and discussion

This final chapter will present the main contributions and implications in Section 8.1. The following section contains a methodological discussion. Section 8.3 presents suggested future research and the final section contains concluding remarks.

7.1 Contributions and implications

The main contribution of this thesis is the developed guidelines;

1. Place the graphical user interface so it is comfortable to explore and interact with
2. Place visual feedback to selections within the immediate interaction area
3. Keep information dense areas interaction-free
4. Use dwell times of various lengths
5. Avoid using time-limited information
6. Never force users to interpret information in movement
7. Use standards and affordances to minimize the cognitive load

They contribute to answering the research question of what aspects should be considered when designing a graphical user interface in virtual reality - controlled by head orientation input - for a qualitative user experience. Furthermore, they lead to an increased understanding of how graphical user interfaces (GUIs), should be designed for the application of video on demand (VoD) services.

Following the expected contributions (previously described in Section 2.8.2), a higher understanding of the different concepts individually has also been achieved since a majority of the guidelines and the analysed results are applicable for other contexts and applications. The results could be relevant for how to use the interaction method of head orientation input in other contexts than virtual reality (VR) and for other applications than VoD services. Furthermore, the guidelines can be applicable for other applications in VR and some even for the use of other interaction methods. Moreover, Accedo has received feedback on their prototype and how to continue the development to improve the UX of the GUI. An unexpected contribution is the suggested term ‘head gaze’ for the interaction method that uses head orientation input to approximate a gaze cursor in the centre of the users’ field of view to use for interaction. Finally, the thesis will contribute as a foundation for future research.

The developed guidelines can be used for future development of VR GUIs controlled by head orientation input or, if rephrased, as heuristics for UX testing of VR GUIs controlled by head orientation input. For the development of VoD services, the result presenting the users pragmatic and hedonic goals including their behaviour, needs, preferences, and context of use could be used as a general persona to develop requirements when designing the UI.

Dorabjee et al. (2015) claim that there has been a lack of focus on the analysis of UIs for HMDs. Furthermore, Bowman et al. (2004) state that there is a lack of empirical evidence for the usability of various system control techniques for 3DUIs. Oculus VR (2015b) claims there are no traditional input methods ideal for VR and that innovation and research are needed. Moreover, the usability of eye gaze as an input modality remains marginal (Bazrafkan et al., 2015).

This research has addressed all those shortages and since the head orientation and head gaze interaction technique does not require any additional input device, other than what is already available in all HMDs for head tracking, the UX guidelines for how to design for and use the technique could increase the use of it. Furthermore, it could increase the use of HMDs. Simple and
inexpensive HMDs such as VR sets used with smartphones could be spread and used for several purposes, such as educational, to a greater extent. If the leading experts knowledge could be shared in a virtual environment to people all over the world, this could have a major impact on the society. Furthermore, this interaction method used to control systems is of specific value for people with disabilities, prohibiting them from using their bodies to interact with systems.

The possible ethical concerns with this research are if the results are used to improve systems for other application of doubtable nature. Possible scenarios for this could be in simulators for military purposes. However, the research and results are primarily considered to have a positive impact on the society.

7.2 Methodological discussion

The research design used a combination of a heuristic evaluation, interviews, usability tests, and a survey and the results from each of the methods were triangulated. Since the research approach has been to mainly gather qualitative data, trustworthiness is used to evaluate it. The trustworthiness of qualitative research is based on credibility, transferability, dependability, and confirmability of the research (Lincoln & Guba, 1985).

Credibility can be established by prolonged engagement, persistent observation, and triangulation (Lincoln & Guba, 1985). The research was prepared, conducted, and analysed over a period of about two months. Limitations in resources and time restrictions prohibited a longer and more comprehensive research. More time to research the different concepts and to validate the results would have been desirable in order to achieve higher credibility of the result. However, both the interviews and the usability tests were sound and video recorded so the analysis of the data could be detailed and comprehensive. Moreover, the credibility can be considered as improved since data triangulation was used in the research.

Transferability refers to the external validity of the result and whether the findings from the research could be applicable in other contexts (Lincoln & Guba, 1985). The study and context for the study have been described in a detailed and exhaustive way. It is the author’s belief that the findings from the research could be applied in other contexts but only people who would like to use the research results could truly validate the transferability of the result.

Dependability can be achieved if the findings are consistent and can be repeated (Lincoln & Guba, 1985). However, people’s opinions and knowledge always change and therefore, the findings could never be consistent even if the research was repeated with the same participants. In order to achieve dependability, the methods, and the methodological decisions have been described in detail. However, it is up to the receiver to decide if the research seems dependable or not.

Confirmability refers to whether a neutral approach has been followed, where the findings are based on the respondents and not the researcher’s bias or interest (Lincoln & Guba, 1985). This has been achieved in this thesis since the researcher has been aware of and described how this was avoided during the procedure of the research methods. Furthermore, both the steps of the research methods and the collected data have been described in detail in order to show the origins of the results in a transparent way.

During the research, there were decisions made and problems discovered that might have had a negative impact on the results. The formulation of one of the survey questions was not descriptive enough which led to it being interpreted in various ways. When this was discovered, a decision was made to exclude the result of the question from the analysis.
During the usability tests, the sound from the prototype was led to the TV in order for the researcher to hear what the user was hearing and this could have affected the UX negatively since the most immersive experience would have been for the user to use earplugs. Additionally, the smartphone in the HMD was connected with a cable to the TV, which might have made the HMD feel heavier and less free for the user to control. This might even have caused the users to feel more limited in their head movements during the test. As a consequence, the analysed data used to develop the first guideline, might not be reliable and the guideline might not be as important as previously stated. To confirm the importance of this, and the other suggested guidelines, more research is required.

7.3 Future research

This research has focused on the UX of head gaze controlled VRGUIs. Research limitations have included time restrictions and resources available. The first suggestion for future research is therefore to expand the research. This could include using additional research methods, additional users in the tests, additional prototypes or systems that are tested, and additional HMDs that are used for the testing. The quality of the data could be improved with the use of more advanced technology, other settings for the tests, and more experienced researchers.

Secondly, a natural step in continuing the research in this thesis would be to validate and refine the developed guidelines. This could include developing and testing prototypes based on the suggested guidelines. Further questions remaining to be answered include if these guidelines are valid across different HMDs, for different applications and for people of all ages, knowledge and previous experience. One of the predicted most common use context is when the user is laying down in bed and for that specific context, more research is needed. Moreover, the research has only tested people living in Sweden with a certain culture. There might be large cultural differences to take into account.

Further suggestions include expanding the research to other applications than VoD services to gain additional insights. Moreover, the usage of HMDs should be observed in a natural environment to gain additional contextual insights.

7.4 Concluding remarks

This thesis has started to answer the research question of what aspects should be considered when designing a graphical user interface in virtual reality - controlled by head orientation input - for a qualitative user experience. Oculus VR (2015b) claims there is no traditional input method ideal for VR and that innovation and research are needed. Even though this thesis contributes to a small part of that research, there is still an abundant amount of research to be done.

When this research started, several HMDs for consumers were just about to be released, and now several of them are released. There has been a surge in media addressing everything about HMDs and VR for consumers and it has been impossible to incorporate all the new information in this thesis. If the thesis research had been carried out today, large parts of it would probably have been different based on the increased information and resources available. With the predicted continuing growth in users of and applications for VR, together with the increase of HMDs on the consumer market, the need for VRGUIs that offers qualitative UX will increase. This will hopefully push the research to continue along the path of this thesis research and extend it even further to address VRGUIs for all kinds of users, applications, and HMDs.
References


Oculus VR (2015a) User interface guidelines. Mobile development documentation, version 1.0.0. Available at Internet:


Appendix A: Heuristic evaluation inspection questions

Heuristic evaluation inspection questions (translated to English - original in Swedish)

1. Is it clear to the user what is happening in the system?
2. Do the system use the users language in a natural and logical order?
3. Is it easy for the user to back or undo an action?
4. Is the system consistent?
5. Is the system preventing possible ‘wrong’ actions?
6. Is the system helping the user so s/he does not have to keep information, interaction possibilities and choices in his or her memory?
7. Is the system adapted for both novice and experienced users so they can use it effectively?
8. Do the system display more than relevant information in such a way that it makes it hard for the user to take in the relevant information?
9. Are error messages designed so the user understand what the problem is and how s/he can solve it?
10. Is it easy to find the help function and does it focus on the tasks the user want to perform? Is it easy to overview and does it list concrete steps?
Appendix B: Interview questions

**Interview questions (translated to English - original in Swedish)**

1. How old are you?
2. How long have you been using video on demand services?
3. Which video on demand services have you used?
4. Which kind of technology do you use when you use video on demand services today?
5. In which situations do you usually use video on demand services?
6. Where do you usually use video on demand services?
7. Are you usually using video on demand services alone or in company with other people?
8. How do you most often decide what you would like to watch?
9. When you watch a video, what do you base your choice of platform and technology on?
10. What functionality do you want to have in a video on demand service?
    a) Is there something you feel like you are missing today?
11. How does the surroundings and environment affect you when you watch a video?
12. How does a typical video on demand usage look for you? Give an example!
13. What aim do you have with using video on demand services? What do you want it to lead to?
14. What, except a good video, do you think create a good video watching experience?
15. When you use a video on demand service, is there any problem or frustration that you often have?
16. What previous experience do you have with virtual reality?
17. How do you feel about watching video in virtual reality?
    a) Is there anything that would change your opinions?
18. How (if at all) would you like to watch video in virtual reality?
19. Where (if at all) do you think you would like to use virtual reality to watch video?
    a) Why?
20. In which situations (if at all) do you think you would use virtual reality to watch video?
    a) Why?
21. Describe ways you think that virtual reality could add to your video watching experience (if any).
22. Describe potential negative aspects you see with watching video in virtual reality (if any).
    a) Do you think these would stop you from watching video in virtual reality?
23. How do you imagine a user interface will look like in virtual reality?
24. Have you ever controlled a system with your eye movements or head movements?
    a) How was the experience?
Appendix C: Questions after usability test

Questions asked to the user after the usability test
(Translated to English - original in Swedish)

1. What kind of feelings did you experience during the use?
   a) Why do you think you were experiencing them?
2. How would you describe this experience compared to your expectations?
3. If you compare watching video with a head-mounted display and how you normally watch video, which do you think give
   the best experience?
   a) Why?
4. What (if anything) would you like to experience with this kind of technology?
5. What do you think is the most interesting and valuable part of watching video in virtual reality?
6. How did you feel about the way you were navigating in the application?
7. Did you feel like you could control everything you wanted to control in the user interface?
8. How did the time it took to make a selection feel like?
9. Was there anything that you felt was hard to understand?
10. Was there anything that you felt was easy to understand?
11. Was there anything that surprised you?
12. Was there any functionality you felt you were missing?
13. Was there anything that annoyed you?
14. What do you think would have made the experience better?
15. What would have made you chose to watch video in virtual reality over watching it on any other type of screen?
16. Was there anything that you liked in the user interface?
   a) Why?
17. Was there anything that you disliked in the user interface?
   a) Why?
18. If you could have changed the user interface so it would become easier to use, what would you have liked to do?
19. If you could have changed the user interface so it would become more fun to use, what would you have liked to do?
20. Of all the things we have discussed during this session, what do you think have been the most important?
Appendix D: Survey design

Enkät om video on demand-användning

Enkätens syfte

Det här enklares syfte är att samla in information om hur personer använder video on demand tjänster (se begreppspartitionering i introduktion) samt informationen kommer att användas i mitt (Sjölund, inlämman) examensarbete som handlade om hur användningsintensitets designas för video on demand tjänster i virtuella miljöer för att skapa två användningsutslaggar.

För att delta behöver du uppfylla följande krav:

- Vara 20-45 år gammal
- Ha använt eller varit aktiv användare av en eller flera video on demand tjänster

Begreppspartitionering

- Video = Alles sorts videos (t.ex. filmer, ansvarsfria och dokumentära).

Villkor

Genom att svara på den här enkäten godkänner du att jag samlar in informationen som du lämnar och att jag får publicera den i mitt examensarbete som är offentlig handling. Du får säkerhets- och medverkan är självständigt fri.

Tack för dit medverkan

Sofie Hijlman

Så often/oftast startar jag en video on demand tjänst för att...

<table>
<thead>
<tr>
<th>Sällan</th>
<th>Ofta</th>
<th>Ganska sällan</th>
<th>Ibland</th>
<th>Ganska ofta</th>
<th>Ofta</th>
</tr>
</thead>
<tbody>
<tr>
<td>video på grund av det att jag inte förväntade mig att vilja se</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>video på grund av det att jag inte förväntade mig att vilja se</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>har en speciell video och jag inte förväntade mig att vilja se</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>försöker titta på en video som jag inte känt på tidigare</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>kontinuerligt om det finns en pålidelig nyhet</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

Ranka följande alternativ efter vad som är viktigast i en video on demand tjänst:

- Jag tittar på och snabbt kan följa de videos som jag tittar på snabbt och självbestämmer vad jag vill se.
- Jag tittar på videos som jag tittar på snabbt och känner mig övertygad att jag inte kommer att förlora några videos som jag tittar på.
- Jag tittar på videos som jag tittar på snabbt och har större praktik som jag tittar på.

När jag vet exakt vad jag vill titta på så hittar jag ofta videoen genom att...

- Söker efter videoen i en handbok
- använder sökmotorer på videosida
- använder sökmotorer på videosida
- hinner titta på videosida
- tittar på videosida

När jag vet ungefär (men inte exakt) vad jag vill se tittar jag ofta...

- söker efter videoen i en handbok
- använder sökmotorer på videosida
- tittar på videosida
- tittar på videosida
- tittar på videosida

Det når (om något) frustrerar mig när jag använder video on demand tjänster (bortsett från internetuppkopplingsproblem och videoutbud i tjänsten):

- Finns inte några videos
- Det tar för länge att hitta videos
- Videospasserar och inte finns
- videos grupperas oförnuftigt
- videos är inte tillgängliga inom slutmarknad

Denna enkät skapades med hjälp av SurveyMonkey.

https://www.surveymonkey.com
## Appendix E: Heuristic evaluation result

<table>
<thead>
<tr>
<th>Problem description</th>
<th>Problem name</th>
<th>Heuristic(s)</th>
<th>Severity</th>
</tr>
</thead>
<tbody>
<tr>
<td>The user looks at a menu option and starts reading the text explaining the option</td>
<td>GLOBAL: Selection of function occurs before user is finished reading and interpreting the function name and icon</td>
<td>1. Visibility of system status 3. User control and freedom 5. Error prevention 6. Recognition rather than recall</td>
<td>High</td>
</tr>
<tr>
<td>but the selection of the option is completed before the user has the opportunity to</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>finish reading and understand the function he/she just selected/activated.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The user choose “movies” in the main menu but does not see any change in the</td>
<td>GLOBAL Hard to see selection result and where to continue interaction to reach goal</td>
<td>1. Visibility of system status 2. Match between system and the real world 6. Recognition rather than recall</td>
<td>High</td>
</tr>
<tr>
<td>VRGUI as a response to the action and does not know where to continue looking for</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>the movie s/he was looking for and reach the goal of starting it.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The user gaze at the button for leak light in the setting menu but does not see or</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>notice the result of the action because s/he is to focused on the button and the</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>change is to far away from the focus point in the persons FOV.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>When the user has done something by accident (sometimes not even aware of the</td>
<td>GLOBAL No way to undo or go back from faulty or accidental selection</td>
<td>1. Visibility of system status 5. Error prevention 6. Recognition rather than recall</td>
<td>High</td>
</tr>
<tr>
<td>selection/activation) there is no easy way to undo or go back from that selection.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>If a selection of a TV show from the display menu is done by accident - the menu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>is replaced with episodes from that TV show and if the user want to go back to se</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>the display of all the TV shows s/he needs to start over by moving to the main</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>menu and select TV shows again. If the user accidentally selects “Exit VR Experience”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>it will immediately close down the application and the user need to remove the</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>smartphone from the HMD to restart the application.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The user have to scroll horizontally by turning their heads to the left or right to</td>
<td>Time consuming and hard to find a specific episode of TV show in display menu</td>
<td>3. User control and freedom 5. Error prevention 6. Recognition rather than recall 7. Flexibility and efficiency of use</td>
<td>High</td>
</tr>
<tr>
<td>search for the right episode. Text is to small to be read while scrolling and there</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>is no clear difference in imagery between seasons or episodes so the user is</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>forced to stop the scrolling, turn his or her head in to the middle to read and</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>then continue to scroll and guess where to stop next time.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The user need to scroll in the display menu by turning his or her head to the</td>
<td>Non-ergonomic interaction required to reach goal</td>
<td>7. Flexibility and efficiency of use</td>
<td>High</td>
</tr>
<tr>
<td>right or left. The turn is beyond the comfortable degrees making the experience</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>uncomfortable and leaving the neck soar.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>When the video is running, the play/pause button is displaying the play icon,</td>
<td>Play/pause button display wrong icon (bug)</td>
<td>4. Consistency and standards 5. Error prevention</td>
<td>High</td>
</tr>
<tr>
<td>confusing the user of the action that is pause if the user dwell long enough on</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>it.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Example 1: When choosing a new movie or TV show episode to watch from the display</td>
<td>Information required for decisions is not visible when user is interacting with menu</td>
<td>1. Visibility of system status 5. Error prevention 6. Recognition rather than recall</td>
<td>Medium</td>
</tr>
<tr>
<td>menu it might be important to know what is currently running on the screen and</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>then the user have to look up from the display menu and at the screen or to the</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>right to find the video information.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Example 2: When the user is fast forwarding or rewinding the user can not see the</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>timeline for the video so it is not possible to tell how far forward or rewind</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>the user have moved into the video.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The icon used for “Exit VR Experience” and “Rewind 10 seconds” is according to</td>
<td>Icons used for “Exit VR Experience” &amp; “Rewind 10 seconds” look like one</td>
<td>2. Match between system and the real world</td>
<td>Medium</td>
</tr>
<tr>
<td>the users previous experience related to actions to “go back” or undo an action.</td>
<td>usually used for a “go back” or undo action</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The same icon is used for two different actions and before the user gaze over</td>
<td>The same icon is used for both “Exit VR Experience” &amp; “Rewind 10”</td>
<td>1. Visibility of system status 4. Consistency and standards 5. Error prevention</td>
<td>Medium</td>
</tr>
<tr>
<td>them there is no explanatory text separating them from each other more than their</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>different placement in the VRGUI.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>When the user is gazing at most object, s/he will see a orange progress circle</td>
<td>Feedback during selection (when hovering) is not consistent across different</td>
<td>1. Visibility of system status 4. Consistency and standards 5. Error prevention</td>
<td>Medium</td>
</tr>
<tr>
<td>filling to track the progress of the selection. But this circle is not used for</td>
<td>objects the user can interact with</td>
<td></td>
<td></td>
</tr>
<tr>
<td>certain selection/activation tasks such as fast forward and rewind. User waits</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>for the normal feedback and it takes longer to realise when s/he has triggered the desired action.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The icons used and the text explaining each function in the setting menu is</td>
<td>Icons &amp; explaining text for setting menus functions hard to understand</td>
<td>2. Match between system and the real world</td>
<td>Medium</td>
</tr>
<tr>
<td>hard to interpret and does not match with the users previous experience and</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>knowledge.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Help function is missing</td>
<td>Help function is missing.</td>
<td>10. Help and documentation</td>
<td>Medium</td>
</tr>
<tr>
<td>Icons (like the two orange in the middle of the setting menu and the play menu)</td>
<td>GLOBAL Unclear what is possible to interact with and what is not</td>
<td>2. Match between system and the real world</td>
<td>Low</td>
</tr>
<tr>
<td>is not possible to interact with even though they look like they could be. The</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>same goes for the progress bar below the display menu.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Information about state of video missing</td>
<td></td>
<td>1. Visibility of system status</td>
<td>Low</td>
</tr>
</tbody>
</table>