Tampere University

Dionysios Georgios Papadimitriou

# USER EXPERIENCE EVALUATION IN VIRTUAL REALITY

## Conducting an evaluation on multiple characteristics of a Virtual Reality Experience

Faculty of Information and Communication Sciences
Master of Science Thesis
November 2019

# ABSTRACT

Dionysios Georgios Papadimitriou: User Experience in Virtual Reality, Conducting an evaluation on multiple characteristics of a Virtual Reality Experience
Master of Science Thesis
Tampere University
Master's Degree Programme in Human-Computer Interaction
November 2019

---

Virtual Reality applications are today numerous and cover a wide range of interests and tastes. As popularity of Virtual Reality increases, developers in industry are trying to create engrossing and exciting experiences that captivate the interest of users.

User-Experience, a term used in the field of Human-Computer Interaction and Interaction Design, describes multiple characteristics of the experience of a person interacting with a product or a system. Evaluating User-Experience can provide valuable insight to developers and researchers on the thoughts and impressions of the end users in relation to a system. However, little information exists regarding on how to conduct User-Experience evaluations in the context of Virtual Reality. Consecutively, due to the numerous parameters that influence User-Experience in Virtual Reality, conducting and organizing evaluations can be overwhelming and challenging.

The author of this thesis investigated how to conduct a User-Experience evaluation on multiple aspects of a Virtual Reality headset by identifying characteristics of the experience, and the methods that can be used to measure and evaluate them. The data collected was both qualitative and quantitative to cover a wide range of characteristics of the experience. Furthermore, the author applied usability testing, think-aloud protocol, questionnaires and semi-structured interview as methods to observe user behavior and collect information regarding the aspects of the Virtual Reality headset. The testing session described in this study included 14 participants. Data from this study showed that the combination of chosen methods were able to provide adequate information regarding the experience of the users despite encountered difficulties. Additionally, this thesis showcases which methods were used to evaluate specific aspects of the experience and the performance of each method as findings of the study.

Keywords: User-Experience evaluation, user-experience, Virtual Reality, usability, head-mounted displays, usability testing

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

## TABLE OF CONTENTS

# 1) INTRODUCTION

What makes Virtual Reality unique as technology is its immersive ability, i.e., to transfer its users to a location other than the one they are currently in. Virtual Reality users have the opportunity to visit lands formerly unreachable, explore situations from new perspectives and connect with people from all over the world in the same digital space from the comfort of one's home. Owning a Virtual Reality device is becoming commonplace in people's homes due to technological progress, reduced prices and newfound interest in the technology. Numerous different applications of Virtual Reality exist that cover a wide range of interests, such as entertainment in the form of videogames[1], education [Abulrub & Williams 2011; Bell & Folger 1997], and healthcare [Wilson *et al.* 1997; Stone 2011]. However, quantity of applications does not necessarily guarantee quality. Designers and developers must create Virtual Reality products that are not only functional, but also provide a full-filling User-Experience to the users.

User-Experience is a top-level definition for characteristics of a system that the field of Human-Computer Interaction is trying to define. The characteristics can be related to broad meanings and topics [Forlizzi and Battarbee 2004], varying from pragmatic and hedonic, to emotional aspects of the experience. Maguire [2013] exclaims that one element of User-Experience is usability. The author explains that the goal of usability is to pay attention to how users succeed in completing their goals when using a product. However, although some research has been devoted to understanding usability in Virtual Reality, less attention has been paid to the topic of evaluating User-Experience in Virtual Reality.

Developers and researchers interested in crafting captivating experiences for Virtual Reality products would ideally consider evaluating as many aspects of an experience as possible, in a holistic evaluation of the user-experience of the product. However, with the amount of Virtual Reality applications that are available, identifying the aspects of User-Experience that are relevant to each specific case can be a daunting and overwhelming procedure, as the characteristics influencing one experience can be numerous. Similarly, locating and selecting the appropriate combination of methods that can be used to evaluate User-Experience in Virtual Reality introduces another layer of complexity. To assist developers and researchers in this evaluation process, this work investigates **how a holistic User-Experience evaluation can be conducted on a Virtual Reality product.**

---

[1] SteamVR is a platform where users can purchase Virtual Reality video game products: https://store.steampowered.com/vr/

In order to conduct this evaluation, I used a virtual reality headset developed by a company whose name cannot be disclosed within this thesis. Respecting their wishes and interests, this evaluation included aspects of the experience and impressions of users that interacted with that headset. Additionally, various methods were applied to collect, measure and evaluate the opinions of the users. This work provides one example of holistic User-Experience evaluation of Virtual Reality, by capturing multiple facets of the experience of this particular headset and utilizes a set of methods to evaluate the experience. The aspects that were included in this User-Experience evaluation were influenced by time constraints, resources of the researcher and the scale of what was feasible to include within one thesis. Similar limitations are to be found both within the industry itself and the academia. Naturally, as there are numerous Virtual Reality applications for different contexts, the combined aspects that can impact a users' experience might differ. As such, this work is describing the impact of the aspects of this particular Virtual Reality system on the overall User-Experience.

Before the evaluation, I collected information from the available literature regarding the aspects that would be evaluated and the methods that could be used to investigate them. Afterwards, a usability test was planned and performed to expose potential issues that the users might be experiencing when interacting with the headset. To further capture the impressions of users, questionnaires were administered and followed by a post-test interview. Large amounts of data were collected from the experiment and organized in order to provide information related to the experience. Categorizing and analyzing the available data proved to be the most laborious and extensive part of this evaluation. The usability test was successful in exposing issues that users experienced and identifying potential root causes of these issues. However, simultaneously conducting and supervising the test was difficult for one moderator due to the numerous tasks that had to be performed at the same time. Questionnaires in combination with answers from the interview questions were able to provide insight on impressions and thoughts of users regarding the experience. Regardless of encountered difficulties, the joint work of the methods applied in this evaluation were able to measure and effectively expose different aspects of the user experience of the headset and its software.

This thesis is organized as follows: Chapter 2 describes related work on User-Experience and usability, human factors and characteristics of Virtual Reality. Chapter 3 describes in detail the holistic evaluation conducted and how it was organized and executed. Chapter 4 portrays the results from the conducted experiment. Chapter 5 provides an answer to the research question of this study by discussing the aspects of this experience and the methods that were used to evaluate the data alongside observations from the experiment. Finally, Chapter 6 summarizes all of the findings of this study.

## 2) USER EXPERIENCE AND VIRTUAL REALITY

To understand User-Experience and the way it can be evaluated in Virtual Reality we need to discuss a number of concepts in this literature review. I will start by introducing the term User-Experience to familiarize the reader with its meaning and highlight its connection to technology nowadays. Next I introduce the subject of ergonomics, their connection to health and safety, and how they can influence the experience of users when it comes to technology and Virtual Reality. The third subsection will explain what Virtual Reality is, its characteristics, and how Virtual Reality works. Additionally, I will describe head-mounted displays and how they can be used to interact with Virtual Reality, as one is used in this thesis. Afterwards, I will provide examples of Virtual Reality applications to give a picture of the state of the technology, connect the concepts of User-Experience and Usability to Virtual Reality and present a collection of methods that can be used to measure user-experience and usability both inside and outside of Virtual Reality.

### USER EXPERIENCE

We live in an age where technology is omnipresent with systems that aid people complete everyday tasks, while at the same time, these technologies are becoming increasingly complex to use. Reducing this complexity has been the task of the Human-Computer Interaction (HCI) field. HCI is related to human characteristics, such as attention, comprehension, human capability of understanding information, and the interplay of information that happens between a system and a user [Holzinger 2013]. This interplay occurs through what we call a user interface (UI), which comprises a large amount of the overall system and is also the place where the interaction between the user and the system is taking place [Seffah & Taleb 2012]. Furthermore, HCI combines together information science and information technology with behavioral psychology [Johnson 2013].

Human-Centered design (HCD) is a technique of designing products by constantly focusing on the requirements and capabilities of the individuals that are meant to use the product [Norman 2005]. User-centered design, a subgroup of HCD, describes the mindset of designing products that allows users to affect the overall design process through a variety of participatory methods, while focusing at the same time on the actual end users of the system [Abras *et al.* 2004].

Furthermore, one of the ways to simultaneously keep in mind the end users of a system and cater for their needs, is by designing for a decent user-experience (UX). According to the official ISO 9241-210:2010 standard, the definition for UX is: “a person's perceptions and responses that result from the use or anticipated use of a product, system or service” [DIS, 2010]. The mindset of designing for a decent user-experience can be adopted to create basically any kind of product, however a great deal of current research is concentrated on the construction of digital experiences.

Directly evaluating the user-experience of a product can be a strenuous, complex and continuous process but an important element of the overall process can be organizing a usability test. The definition of the term usability has been respectively specified by the ISO 9241–11 as: “the extent to which a product can be used by specified users to achieve specific goals with effectiveness, efficiency and satisfaction in a specified context of use.” [Bevan *et al.*, 2015]. In practice, conducting a usability test is a technique that is used to assess a user interface by observing how users interact with the system in a controlled environment. Through the observation stakeholders that are involved in the creation of the system have the chance to detect issues and pain points that the users might be encountering in order to make changes that can improve the functionality of the system.

## HUMAN FACTORS

According to Salvendy [2012], the phrase human factors (alternatively called ergonomics) describes the scientific discipline which utilizes information that is related to human attributes with the goal of creating and designing systems that are suitable to those characteristics. Additionally, the author describes how ergonomics are constantly striving to enhance human comfort in regard to health, reduce potential risks in the work environment and increase overall productivity and efficiency.

However, as the nature of ergonomics can vary greatly depending on the system at hand, Wilson [2000] describes that ergonomics can also be broken down into smaller, more specialized groups; which include cognitive ergonomics, social ergonomics and physical ergonomics. To highlight the importance of physical ergonomics in the digital era, Wahlström [2005] describes how the rapid increase and usage of visual screens in the workplace has been connected to various musculoskeletal symptoms such as neck, wrist and shoulder pains in users. Commonly, in order to ensure the health and safety of individuals, ergonomic evaluations are used as a way to assess the usability of the system at hand which intrinsically requires paying attention to how users physically interact with a system.

In the context of Virtual Reality nowadays, users often have to wear a head-mounted display (HMD) that is attached to their heads for considerable amounts of time. Similarly, other parts of the human body, such as hands, torso and feet of the users might be required to hold or wear equipment to interact with the virtual reality environment in the form of controllers or haptic suits. As a result, user movement can end up being confined within a certain part of the usage environment due to the cords that are attached to the Virtual Reality equipment and the computer, while at the same time, the users are trying to navigate around cords and other items in the room. The combination of wearing HMDs for large amounts of time, small space for mobility within an area and navigating said space can potentially involve a level of physical risk for users. Sharples *et al.* [2007] stated that not paying attention to the usability of the Virtual Reality control devices and how the devices eventually interact with the interface can heavily impact the Virtual Reality application experience.

## VIRTUAL REALITY

Virtual Reality attempts to produce a realistic 3D setting where a user is given the ability to navigate and interact with the generated scene [Gutierrez *et al.*, 2008]. The author further describes navigation as the opportunity of exploration inside the produced scene, whereas interaction is explained as the capability to choose and control entities of the environment. To put in another way, Virtual Reality is trying to discover solutions which can offer convincing, life-like experiences to our senses [Hale & Stanney 2014]. The individuals who operate the Virtual Reality System get the impression that they are within a virtual 3D setting or virtual environment (VE). Giving the impression of a virtual environment to a user can occur inside a CAVE (short for cave automatic virtual environment), a cubic shaped space that is comprised of walls which can produce images in order to create an immersive environment [Muhanna 2015], or via wearing headsets or glasses that position a screen in front of the view of the user.

Geszten *et al.* [2015] explain that realistic 3D environments are optimal for detailed exploration and planning behaviors of users due to the similarities that can be found between how we interact with the actual world around us and the possibilities offered in a virtual environment. However, the authors explain that 3D environments present additional options to the ones we are capable of achieving in the real-world when it comes to cooperation and managing information. For example, digital environments allow users to cooperate with each other from a distance, manipulate the environment in ways that would not be possible in the real world, and also provide a hypothetically unlimited space for storing information.

Furthermore, to be able to better comprehend how individuals can interact with a virtual environment, it is crucial to discuss the concept of degrees of freedom (DoF) and how they correspond to movement within a virtual environment. Virtual Reality systems can afford from three to six degrees of freedom. Six degrees of freedom refers to the freedom of movement of a rigid body in three dimensional space; a body is free to move forward or backwards, up or down, as well as left and right, combined with rotation of three perpendicular axes, often termed pitch, yaw and roll [Paul 1981]. The combination of these six degrees of freedom can demonstrate movements that are similar to the ones that individuals can do in the real world. Furthermore, there are two options offered to the users in order to explore a six degree of freedom virtual reality setting: a) to provide commands to a virtual reality system via a controller or b) by controlling the environment with movements of their head. Controlling virtual reality environments with the movements of the head has been shown to be the best approach regarding overall usability [Chen *et al.*, 2013], and is also the way that users controlled the virtual environment in this study.

## HEAD-MOUNTED DISPLAY

*Figure 1. A person wearing a head-mounted display. Image courtesy of the researcher.*

Head-mounted displays nowadays have captured notable attention and triggered vast economical attempts in order to advance technology for a widespread scope of applications [Azuma *et al.*, 2001]. Take first, for example, the case of augmented reality. Users can wear a head-mounted display on their head that superimposes information on the surrounding world whilst not obscuring their vision [Feiner 2002]. This projected information has allowed innovations in multiple fields. In the field of medicinal surgery, the application *ARassist* aims to increase the performance of first assistants in robot-assisted laparoscopic surgeries by recreating visual representations of surgical instruments and also providing information related to stereo endoscopies [Qian *et al.*, 2018]. In the field of educational learning, augmented reality head-mounted displays have been used to assist young children in reacting to hazardous situations by projecting digital objects within their view in fictional scenarios [Mitsuhara *et al.*, 2017].

On the other hand, in Virtual Reality, head-mounted displays serve the purpose of entirely concealing the visual field of a user to immerse them in a realistically generated virtual environment or video experience [Hua 2017]. To achieve this isolation effectively, head-mounted displays commonly come in the form of a helmet that is worn directly on the user's head. Moreover, the way head-mounted displays work is by constantly presenting visual imagery to the eyes of the user in close combination with the movements of their heads; what is being shown to the user always remains in a set and relative position to the users eyes and ears [Craig *et al.*, 2009]. A few examples of popular Virtual Reality headsets available on the market today include HTC Vive[2], Oculus Quest[3] and Samsung Gear VR[4].

Individuals in the physical world have multiple ways in which they can interact with the environment around them. We use our hands and legs in order to perform functions such as navigating, walking or grabbing objects, which let us manipulate our surrounding environment when needed. We can use our perception of vision to observe and make sense of the world around us in combination with the sense of smell and touch. In Virtual Reality, researchers have been exploring and appraising solutions that allow us to interact and manipulate objects in 3D environments [Mine 1995; Poupyrev *et al.*, 1997; Soukoreff *et al.*, 2004; Teather & Stuerzlinger 2013]. Various interaction techniques that allow users to select objects have appeared, such as having virtual palms [Poupyrev *et al.*, 1996], pointing with a beam [Mine 1995], Leap motion [Khademi *et al.,* 2014] and image plane interaction [Pierce et al., 1997].

Moreover, latest head-mounted displays operate in multiple ways in order to interact with the virtual environment. The HTC Vive and Oculus Quest headsets utilize controllers to receive input from the users in combination with the aforementioned interaction techniques, whilst other headsets, such as Samsung Gear VR and Hololens[5], operate by utilizing the movements of the user's head [Qian & Teather 2017]. Hololens additionally utilizes hand gestures alongside head movements. Blattgerste *et al.* [2018] state that headsets that primarily operate through head movement allow the users to point to their option to select it. This type of selection is made by hovering a digital cursor in the center of the display above the option that the user wants. Consecutively, when the cursor and desired option meet, the selection is counted as successful and the option is then chosen. Another form of selection without utilizing controllers in head-mounted displays can be done by allowing selections using the eyes of participants which is also known as eye-tracking [Qian & Teather 2017]. The selection can occur at different speeds depending on how long the user has been focusing on a particular object, also known as dwelling time [Piumsomboon *et al.,* 2017].

Performing head-movement selections in Virtual Reality has been shown to introduce lower

[2] For more information on HTC Vive please visit: https://www.vive.com/eu/product/
[3] For more information on Oculus Quest visit: https://www.oculus.com/quest/?locale=en_US
[4] For more information please visit: https://www.samsung.com/global/galaxy/gear-vr/
[5] For more information on HoloLens please visit: https://www.microsoft.com/en-us/hololens

error rates in user inputs while simultaneously being appealing to users in comparison to eye-gazing [Minakata *et al.*, 2019]. Hansen *et al.* [2018] performed an experiment where they showed that head-selection accuracy came second to mouse selection in Virtual Reality but with significantly better results in terms of selection errors and overall throughput than eye-tracking. Alas, both aforementioned types of interaction are becoming progressively appealing amongst developers such as Google and Apple due to the advantages of interacting with a virtual environment without using hands or controllers [Jalaliniya *et al.*, 2015].

## APPLICATIONS OF VIRTUAL REALITY

Virtual Reality technology provides opportunities not only for entertainment, but for a variety of fields including teaching, medicine and business. In the field of commerce, Virtual Reality is fitting for digital cooperation and development of products [Kan *et al.*, 2001] along with merchandising and brand promotion [Van Kerrebroeck *et al.*, 2017]. Similarly, Virtual Reality can be an ideal approach for the vacation industries, with implementations in administration, advertising products, amusement, familiarization of individuals with certain locations and conservation of culture and pedagogy [Guttentag 2010].

Virtual Reality can also be an educational tool in other domains, particularly in the field of manufacturing and science in general [Schofield 2012]. It also serves as a sublime space for replicating different kinds of teaching scenarios when needed - be that in military activities [Siu *et al.,* 2016], professions that require handling of hazardous materials [Haase & Termath 2015], or coaching people with ASD (autism spectrum disorder) in the interviewing processes required for work employment [Smith *et al.,* 2015].

In the field of medical science virtual reality presents advantages to both specialists and ailing individuals in various ways; from teaching new doctors in how to perform surgery and handling pain of patients, to healing of psychological disorders [Li *et al.,* 2017]. It has been applied as a curative instrument in emotional and bodily recovery [Cherniack 2011; Cao 2016] and has proven to be a very effective tool in assisting with treatment of mental issues in psychology, such as Schizophrenia [Uvais 2015], overcoming personal fears [Stanica *et al.,* 2016], and Dementia [Garcia-Betances *et al.,* 2015].

However, one of the most popular mediums of using Virtual Reality today is through entertainment and video games. As stated by Hamari and Keronen [2017], throughout the past ten years video games have become a prominent form of amusement, social identity and a commonplace element in people's lives. Furthermore, 59% of the overall populace of the United States is engaging in video game activities whilst the profits of the video game business is surpassing the amount of 15 billion in US dollars [ESA, 2014]. Technological giants such as Facebook, Google and Valve have all indicated their individual interest in Virtual Reality technologies. Valve corporation, a game developer company with numerous critically acclaimed video game titles that also created the digital software platform Steam, announced in 2015 their collaboration with the company HTC in order to create the HTC Vive virtual reality headset [Kim 2016]. Valve has also created their personal Virtual Reality hardware called Valve Index[6]. Similarly, Facebook purchased the American company Oculus, which specializes in virtual reality software and hardware, for 2.3 billion dollars in March 25th of 2014 [Plunkett 2014], whereas Google created the Google cardboard system[7] which aims to inspire attention and growth in the development of Virtual Reality applications.

## USER-EXPERIENCE AND USABILITY IN VIRTUAL REALITY

Donald Norman referred to the phrase user experience as every faucet of an experience of an individual when interacting with a system back in the 1990's [Norman *et al.,* 1995]. UX additionally depends on various movements [Rogers 2012], such as affective design [Joran 2012; Norman 2014], activity theory [Kuutti 1996] and usability research [Nielsen 1994]. Lallemand *et al.* [2015] points out that usability research in the 1980s served as the foundation for the subject of UX as a whole, with the quality of user-experience depending on usability, as applications and software increased in complexity. The authors then continue that in leading UX models usability issues related to effectiveness and efficiency were subsumed as part of the "instrumental" properties of a product. Similarly, Sharp *et al.* [2007] highlights that usability is critical to UX and that various features of UX are connected to the usability that a product has.

Nowadays, to ensure effective usability in conventional computer programs we depend on thoroughly investigated and tested methods and heuristic approaches [Stanney *et al.*, 2003]. In comparison, usability in Virtual Reality is still at an early stage in locating and organizing usability characteristics [Gabbard 1997; Kalawsky 1999]. However, although relatively little research on Virtual Environment usability has been carried out, there have been prior works to refine the experience of virtual environments by standardizing a way of creating virtual environments and conducting usability assessments [Bowman 1999; Gabbard 1997; Kalawsky

[6] For more information on Valve Index please see: https://store.steampowered.com/valveindex
[7] For more information on Google cardboard see: https://vr.google.com/cardboard/

1999].

One of these attempts is the the Multi-criteria Assessment of Usability for Virtual Environments (MAUVE), a taxonomy of criteria and Virtual Environment heuristics that aims to produce an organized way of attaining effective usability and user experience when creating virtual reality experiences [Stanney *et al.*, 2003]. Criterias in the MAUVE system include: a) wayfinding, b) navigation, c) object selection and manipulation, d) visual output, e) auditory output, f) haptic output, g) simulator-sickness, h) engagement, i) presence, j) immersion. In this subchapter I will be expanding on wayfinding, object selection and manipulation, visual output, comfort, and simulator-sickness, presence and immersion due to them being relevant to this research. Thus, the criteria serve this research by providing information to topics that can influence usability in Virtual Reality and provide insight into topics that can be relevant to user experience and usability evaluations of virtual reality environments.

**Wayfinding**

According to Darken and Sibert [1993], wayfinding is the capability of an individual to understand their position and direction whilst traversing inside a manufactured space. Moreover, finding one's way inside a virtual environment is achieved when users maneuver their point of view in order to navigate their surroundings more effectively [Bowman 1999; Bowman *et al.*, 2002]. It has been noted that users in virtual reality often have trouble retaining information related to their current position and direction while navigating [Chen and Stanney, 1999; Darken and Sibert, 1993; Ellis and Meyer-Arendt, 1992]. Naturally, a virtual environment should provide the users with solutions to overcome this. Furnas [1997] suggests that a user should be capable by what they see to strategize the nearest path to reach their objective; similarly, Wickens and Baker [1995] state that the virtual environment should bear enough clues to allow the user to know where they can go and how they can achieve that. Lastly, there should be adequate information offered to the user that explains the nature of the surrounding environment and how the user can recognize and discover objects of choice so that conflicts with identifying their current position and where they want to go can be avoided [Stanney *et al.*, 2003].

**Object selection and manipulation**

The second criteria in MAUVE is the ability to select and manipulate objects inside a Virtual Environment. Similarly to traditional graphical user interfaces, comprehending, selecting and optimizing the ways in which users utilize and interact with items in virtual environments and solve respective challenges, is a crucial part of creating successful virtual experiences [Poupyrev 2000].

Examples of interaction techniques in Virtual Reality vary in nature and include methods to select objects by casting a beam, virtual hands or the ability of users to directly engage with items. Bowman *et al.* [2000] highlighted that one of the most crucial characteristics when it comes to interaction in virtual environments is how consistent the interactions that occur in the real world are in relation to their digital counterparts. Precisely detecting the movements of users in Virtual Reality is paramount to effective usability [Stanney *et al.*, 2003], as mistakes in location and direction can result dissimilarity in movement between the physical and digital environment [Kalawsky 1993]. Another criteria to consider in regards to manipulating objects is that users should be able to choose, alter the location, and change the characteristics of items (such as its hue and form) in the virtual environment; similarly, the affordance of selection of virtual items should be always evident and unambiguous [Kalawsky 1999]. Lastly, the devices that allow users to select 3D objects in Virtual Environments should have very small latency and fast response rate (<50ms), to ensure a smooth and natural object selection for the user [Ware and Balakrishnan 1994].

**Visual output**

The third criteria to consider for usability in virtual reality is the visual output characteristics of the system. To observe the world around, we rely on our sense of vision and hearing to observe the cues around us. In this study, the audio aspect of the experience was not considered as the participants listened to the audio of the experience from a laptop in order to be able to hear the researcher. When it comes to computer interfaces, it has been shown that optically displayed images and hints can be captivating to the users [Kalawsky 1993]. Therefore, in virtual reality, the graphical user interface can give important data related to the digital surroundings of the users, with the goal of making the world more comprehensive [Durlach and Mavor 1996]. However, trying to make the best out of our sense of vision to achieve certain outcomes (e.g. directing gaze of viewers, utilizing white spaces) often presents a formidable challenge to the visual creators of the digital environment, which in turn can also affect the overall usability of the system [Stanney *et al.*, 2003]. One of the reasons is because users are adept at detecting even the smallest inconsistencies of what is being exhibited, such as graphical inconsistencies or stuttering from the visual content [Kalawsky 1999]. As a result, even when visual cues are applied and used in the virtual environment effectively, the attention of users can end up being distracted by the technical aspects of the virtual reality system. Gabbard [1997] state that along with information presented to the users by the designers, we should be taking into account also technological characteristics of the system such as: “stereoscopic support, spatial resolution, field-of-view, update rates, refresh rates, and user comfort and acceptance”. Respectively, the image of the virtual environment being shown to the users should have smooth frame rates (aiming for above 90 frames per second) and small latency [Richard *et al.*, 1996; Ware and Balakrishnan, 1994]. Pausch *et al.* [1993] state additionally that the field of view (what the users can see in the virtual environment including their peripheral vision) and what is being shown to the user should also be corresponding with their head movements. Lastly, Kalawsky [1993] explains that although a 100-degree angle of field of view is needed to replicate the feeling of a realistic virtual environment, it is also additionally crucial to carefully consider what FOV is needed to perform necessary actions in the respective system.

**Comfort**

The fourth criteria is related to the comfort of the user when using Virtual Reality. McCauley-Bell [2002] states that ensuring overall pleasantness and well-being when users interact with Virtual Reality equipment is of utmost priority. Durlach and Mavor [1996] also highlight the importance of users feeling comfortable when wearing head-mounted displays, as potential physical pain or unpleasantness can create negative thoughts and emotions regarding the experience. Discomfort in virtual reality can be the outcome of both physical and visual characteristics of the system, and can affect the usability and overall experience. Physical discomfort in virtual reality may occur due to very tense head-mounted display straps or from weighty and movement-restraining apparatuses [Stanney *et al.*, 2003]. As such, extended usage of interaction devices that allow users to control the virtual environment should not result in overall tiredness to the users [Card *et al*., 1991; Gabbard 1997; Zhai 1995]. Visual discomfort can be the result of issues such as: deficient contrast; badly illuminated environmental spaces, and conflicting depth perceptions that can disorient users when they try to focus on objects of the virtual environment [Rushton *et al*., 1994; Wann and Mon-Williams, 2002]. Additionally, users should not experience extreme eye fatigue whilst interacting with a virtual reality system [Kalawsky 1999]. To summarize, devices that allow users to interact with the environment via hands or other parts of the body, should be ease to use and should not lead to pain due to excessive work of big muscle groups [Hannaford and Venema 1995; Zhai et al., 1996].

**Simulator-Sickness**

Another criteria which can affect overall system usability and impact user experience is Virtual Reality sickness symptoms. In general, there have been concerns regarding the on-going evolution of Virtual Reality technology due to the existence of motion-sickness effects that affect a noteworthy amount of virtual reality users [Chien and Jenkins, 1994; Stanney *et al.*, 1998]. Stanney *et al.* [2003] explain that there is a significant neurological conflict that occurs because of what the users are comprehending and what they are expecting from their environment to happen in the moment. The authors explain that this happens as the users are aware that they are concurrently inside a virtual world and physical world while using Virtual Reality. This neurological clash is considered to be the principal culprit of motion sickness in Virtual Reality [DiZio and Lackner 1992]. In order to avoid possible sensory conflicts that create the feeling of motion sickness, users predominantly quit their usage of a Virtual Reality system or adjust to their physical environment [Reason and Brand, 1975]. The authors continue to explain that users might continue to feel symptoms of sickness whilst readjusting to their normal environment, even after they have stopped using the virtual system. Ideally, users should not experience symptoms such as discomfort, headaches, eye-strain, salivation, sweating, nausea and burping when interacting with a Virtual Reality system [Kennedy *et al.*, 1993]. The overall user-experience Virtual Reality will be increased as we continue to study and comprehend the reasons behind motion-induced sickness in virtual environments and how we can design our virtual experiences in order to not evoke such responses. Recently, Oculus has released a list of readings[8] and best practises[9] related to cybersickness.

**Presence and Immersion**

As stated by Steuer [1992], presence is the predominant objective of Virtual Reality and the characteristic that describes the nature of the medium best. Even though there has not been a complete agreement on the exact definition of the term presence, it can be described as the personal psychological impression of the user that the Virtual Environment is their natural environment, instead of their physical one [Stanney and Salvendy 1998]. In other words, not only will users that are deeply present in Virtual Reality consider the surrounding environment as more alluring, but additionally regard the environment as a location that they have actually been and not as generated content [Aitamurto *et al.,* 2018]. Riva *et al.* [2014] additionally explain that participants who experience higher presence when performing an action will not only feel more engaged during the activity, but potentially be more successful in what they were trying to achieve (e.g. what they intended to do and what happened). Similarly, the level of presence that an individual can experience inside a Virtual Reality system can have an impact on the usability of the system itself [Fontaine 1992; Zeltzer 1992]. As such, developers of Virtual Reality experiences should take into account factors that can impact and influence the sensation of presence when designing or evaluating their Virtual Reality systems. Such factors can be the implementation of 3D immersive sound, how well the system responds to the inputs of users, and the ability to navigate inside the

[8] The readings can be found here: https://developer.oculus.com/design/latest/concepts/bp-reading/
[9] Best practises founded here: https://pdfs.semanticscholar.org/0d8b/1d9d32bebdc79143f07ad673d97dac230cfe.pdf

Virtual Reality experience [Cummings and Bailenson 2016; Slater and Wilbur 1997].

Another vital component that may lead to greater sense of presence is the concept of immersion [Slater and Usoh 1994; Slater *et al.* 1994; Witmer and Singer 1998]. Contrary to the psychological sense of presence, immersion can be described as the compilation of technological components of a system that serves the purpose of achieving an engrossing, captivating and evocative representation of reality to the users [Aitamurto *et al.* 2018]. These technological features include the field of view of the users, accurate detection of body movements, framerates, system latency, overall visual and hearing quality and the authenticity of the environment [Bowman and McMahan 2007; Cummings and Bailenson 2016; Slater and Wilbur 1997]. Slater and Wilbur [1997] explain that we can consider a virtual reality system to be more immersive than another one when the first system is surpassing the latter in at least one of the aforementioned technological features. Similarly to presence, immersion has been considered as a crucial factor that contributes to the usability of a Virtual Environment [Stanney *et al.*, 2003].

Imagine User-Experience in Virtual Reality as a puzzle, where the multiple individual usability characteristics of Virtual Reality (e.g. simulation sickness, interaction techniques, wayfinding, comfort ergonomics, presence) are the jigsaw pieces that when put together, help create the puzzle itself. User-Experience in Virtual Reality encompasses multiple aspects of the experience that a user is having while inside a Virtual Environment, and is consecutively influenced by these multiple characteristics of the experience at the same time. To further support the notion of this idea, the following figure 2 depicts the usability criteria mentioned in this chapter as part of User-Experience in Virtual Reality. The figure could serve as a guide to Virtual Reality aspects for developers to understand User-Experience in Virtual Reality. As not every aspect related to Virtual Reality is described in this thesis, the graph could be enriched with the addition of other criteria.

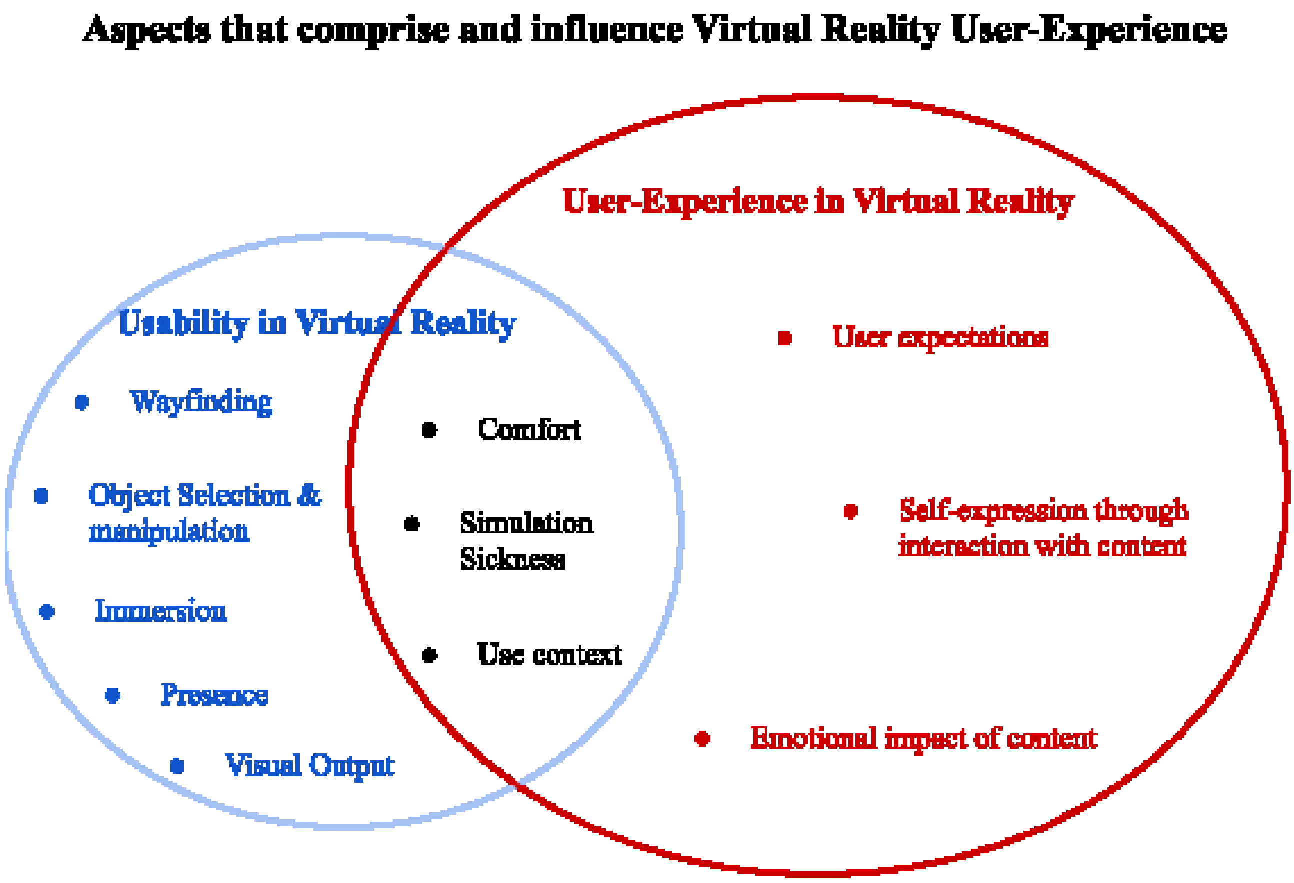


*Figure 2. Various Aspects that can influence the experience of the user in Virtual Reality.*

## MEASURING USER EXPERIENCE IN VR

As mentioned, ISO 9241-210:2010 standard provides a definition to the term User-Experience. However, despite the existence of the ISO standard, the term User-Experience still remains a convoluted and complex term that is continuously being discussed in both the industry and academia [Obrist *et al.*, 2009; Roto *et al.*, 2009].

Defining the metrics to assess User-Experience of a product can be challenging as User-Experience can be very subjective and vary greatly depending on how the user feels at the moment and the current context [Alben 1996; Hassenzahl and Tractinsky 2006; Mäkelä and Fulton Suri 2001]. To avoid this, Roto *et al.* [2009] suggest creating metrics that are based on the characteristics of the product itself instead of the users, as products are already created with a specific user experience in mind (e.g. being enjoyable, safe or calming to the users). This idea is similar to the notion of Experience-Driven design, where designers should evoke a certain experience through their product rather than creating an experience in itself [Sanders and Dandavate 1999; Wright *et al.,* 2005]. While trying to create a product with characteristics that can possibly convey a certain experience to users, designers come up with User-Experience goals which help them keep the experience of the product in the heart of the overall product creation process [Kaasinen *et al.*, 2015]. Roto *et al.* [2009] state that the more definite the product, the more solid User-Experience metrics or Goals can be, which then allows certain methodologies to be chosen to gather data and assess these metrics.

User-Experience can have both a functional and an emotional side to it [Hassenzahl 2003; Jordan 2002; Norman 2004]. The decision of the methodologies should directly correlate to the type of metric that we are concerned about, whether pragmatic or hedonic, as there are multiple possible criteria for different methods [Väänänen-Vainio-Mattila *et al.* 2008]. Since User-experience is relying on the context, it would be impossible to utilize only just one method to measure User-Experience as a whole. As such, we should familiarize with available tools and methods that can be used depending on the context [Väänänen-Vainio-Mattila *et al.* 2008]. There are many known techniques that can be used to assess the usability of a product, as research on usability has been organized exhaustively over the years (e.g., Law *et al.* [2009], and various websites[10]). Similarly, Vermeeren *et al.* [2009] mention a list of 96 collected User-Experience evaluation methods that are accessible online[11].

Finding the right approach to measure User-Experience in the context of Virtual Reality can be challenging and complicated. One of the main issues is the lack of standard ways for researchers to assess and evaluate virtual environments and other emerging technologies [Winn 2002; Neale and Nichols 2001]. Bowman *et al.* [2002] explain that traditionally established methods for evaluating software applications may prove to be unfitting and inefficient when it comes to Virtual Reality, due to the unique ways that users can interact with the content. Another challenge is that the experiences of the users can vary to a substantial degree even in controlled experiments [Neale and Nichols 2001]. The authors explain that this occurrence has been observed while studying how Virtual Reality can affect the way people felt, positively or negatively, after a certain amount of time. It was shown that the experience of presence, pleasure, cybersickness differs from person to person [Nichols 1999]. The diverse range of Virtual Reality applications currently existing and being developed on a number of different fields (e.g. tourism, medicine, education, business), is only adding to the overall complexity.

For researchers to capture aspects of this complexity, there are multiple questionnaires that assess experiences of users, such as general reactions [Nichols 1998, 1999; D'Cruz 1999], user

[10] http://www.usabilityhome.com/, http://www.usabilityfirst.com/glossary/cat_66.txt
[11] http://www.allaboutux.org/all-methods

incentive [D'Cruz 1999], presence [Witmer and Singer 1995; Slater *et al.*, 1994; Psotka and Davison 1993; Nichols *et al.*, 2000], entertainment [Nichols et al., 1997; Nichols 1999] and cybersickness [Kennedy *et al.,* 1993; Kolasinski 1996; Regan 1994; Nichols et al., 1997; Cobb et al., 1999]. Additionally, numerous studies on Virtual Reality describe context specific aspects of Virtual Reality experiences. For example, Steven [2016] provides topics such as problems, symptoms, and general advice related to VR with the goal of informing developers and engineers in their efforts to create a refined experience in Virtual Environments. Similarly, Mine [1995] explains the ways users can interact with the virtual world in the form of interaction techniques, with the goal of providing the reader with an understanding of how to develop a realistic and innate graphic user interfaces in Virtual Environments. As the academic literature on how to conduct and evaluate User-Experience practises is sparse at best when it comes to Virtual Reality applications, it could prove beneficial to try and study as many of the related aspects of the experience as possible, in order to make adequate inspections and measurements on the context. However, it is out of the scope of this thesis to provide a detailed descriptive overview of all the possible relevant topics that are related to Virtual Reality applications.

## MEASURING USABILITY IN VR

Montero *et al.* [2008] state that despite the challenges that are entailed when evaluating personal opinions of users interacting with a product, usability evaluation should always be conducted at the same time. While it has been shown that Virtual Environments can be plagued by crucial usability problems, such as users getting confused by the environment or having difficulties when interacting with items [Kaur *et al.,* 1996], there has been a growing amount of research investigating and exercising usability in Virtual Reality [Gabbard *et al.*, 1999; Deol *et al.*, 2000; Johnson 1999]. As such, a brief descriptive list of usability evaluation methods that can be applied in Virtual Reality follows herein.

First and foremost, Bowman *et al.* [2002] assembled an index of usability evaluation practises that were originally created for traditional graphic user interfaces but were later adapted to Virtual Reality usability evaluations. The methods included in their index are: a) Cognitive Walkthrough, b) Formative Evaluation, c) Heuristic Evaluation, d) Post-hoc Questionnaire, e) Interview/demo and f) Summative or Comparative Evaluation.

The primary goal of the *Cognitive Walkthrough* method is to comprehend the usability of a user interface in regards to users that are getting familiarized with it for the first time; as such, the evaluation observes how well the interface is able to assist the user on standard, everyday tasks (see, Polson *et al.* [1992]). The method of *Formative evaluation* (see, Scriven [1967], and Hix and Hartson [1993]) tries to assess via observation how well a user interacts with a system by making the user perform specific tasks with the goal of exposing usability problems, and secondly by assessing how well the system assists the user in getting familiarized with the system (e.g. the system environment, the outcome of specific tasks and overall user learning). Consecutively, depending on how formal or informal the nature of the experiment is, the formative evaluation can produce both quantitative and qualitative results. *Heuristic evaluation* (see, Nielsen and Mack [1994b]) combines the knowledge of usability professionals to assess an interface by utilizing a list of design rules; afterwards, observations that are collected from the professionals are merged and categorized to fix usability problems that were identified. The method of *Interview/Demo* (see, Hix and Hartson [1993]) predominantly gathers data from users via conversation to get a deeper understanding of their personal opinions in regards to the evaluated interface. Depending on the nature of the interview, there might be a predetermined list of questions being asked or have a more relaxed, improvisational style where the interviewer after asking questions can be more flexible about the topic being discussed. Occasionally, during the interview, the participants might be presented with a prototype version of the product to assist the participants in describing their experience. The *Summative or Comparative Evaluation* method (see, Scriven [1967], and Hix and Hartson [1993]) creates a statistical juxtaposition between two or more versions of a product at hand. The method can, for example, compare the design aspects between the two interfaces, parts of the interface that comprise the interface itself, or solutions that users found to a particular problem. Similar to the formative evaluation, selected participants complete given tasks while the moderators collect qualitative and quantitative data.

Other researchers have been focusing on methods that can be used to evaluate aspects of usability in Virtual Reality. The method of *Testbed evaluation* has been used as a tool to evaluate the performance of Virtual Reality interaction techniques by utilizing design frameworks and numerous variables (e.g. users, tasks, environments, performance metrics), in order to provide a holistic portrayal of performance characteristics of interaction techniques [Bowman *et al.,* 2001; Bowman *et al.*, 2002]. Bowman *et al.* [2002] additionally describe the *Sequential approach* as a method that tries to refine the graphical user interface of a Virtual Reality environment by combining the results of different evaluation techniques (e.g. user task analysis, heuristic evaluation, formative and summative evaluation) in a sequence, while trying to be economically effective. Furthermore, *Theme-based content analysis* (TBCA) is a flexible qualitative method that can expose usability problems by analyzing and categorizing unprocessed data in a coherent, orderly fashion. The unprocessed data is typically collected through a combination of data collection methods (e.g. observation, questionnaires, interviews) that depend on the Virtual Reality context. The goal is to present the analyzed data in a meaningful and comprehensive way to the developers and designers of the Virtual Environment [Neale and Nichols 2001]. Lastly, Gabbard *et al.* [1999] created a *user-centered design approach* that strives to improve usability in Virtual Environments by consecutively using a combination of existing usability methods (e.g. user task analysis, heuristic evaluation, formative and summative evaluation) in an iterative fashion to produce results that enhance the way that users interact with the Virtual Environment.

It could prove useful to start collecting the existing usability evaluation methods related to Virtual Reality together in one place, presented in a coherent and compact manner. Table 1 is a list of existing methods alongside their original source, by building on the paradigm of Bowman *et al.* [2002]. The table is comprised of two columns that list usability evaluation methods and the example of the method in practice. As research in Virtual Reality continues to grow, the list of existing methods can be further enriched and complemented by the newly discovered methods to provide easy access to people interested in usability methods that can be applied to Virtual Reality.

| **List of Usability evaluation methods applicable in VR:** | **Example of the method(s) in practice:** |
|---|---|
| Post-hoc questionnaire | Slater *et al.* [1995] |
| Summative evaluation, post-hoc questionnaire | Darken and Sibert [1996] |
| Informal summative | Bowman and Hodges [1997] |
| Heuristic evaluation, cognitive walkthrough | Steed and Tromp [1998] |
| Formal summative, Interview | Bowman *et al.* [1999] |
| User-task analysis, heuristic evaluation | Gabbard *et al.* [1999] |
| Formative evaluation, summative evaluation | Hix *et al.* [1999] |
| User-centered design and Evaluation | Gabbard *et al.* [1999] |
| User task analysis, heuristic evaluation, formative evaluation | Stanney and Reeves [2000] |
| Theme-based content analysis (TBCA) | Neale and Nichols [2001] |
| Testbed evaluation, sequential evaluation | Bowman *et al.* [2002] |

*Table 1. A list of Usability Evaluation methods applicable in Virtual Reality.*

# 3) HOLISTIC EVALUATION OF A VR SYSTEM

## Conceptualizing the idea of this thesis

This thesis was done in collaboration with a company whose name needs to remain unnamed per their request. The company is in the Virtual Reality market and working on developing a software application on an available Virtual Reality headset. As such, their main interest was to observe how well their software functionality is perceived by the common user. This created the need for understanding and evaluating the experience of the user while they interact with the software of the headset. Thus, it was decided that the focus of the thesis would be to observe and evaluate how users interact with the experience and provide information.

## Research question & goals of evaluation

To take into account multiple aspects of the experience as well as the impressions of users in this thesis, a holistic user-experience evaluation was deemed to be the way to proceed. The term "holistic" in this case is referring to aspects of the experience that can be summarized as users' impressions from the tutorial, interaction with the user-interface and the overall comfort, satisfaction and sense of presence felt throughout their interaction with the headset and its software. The research question of this thesis is: **How can we perform a holistic user-experience evaluation of a Virtual Reality headset?** A holistic evaluation would require evaluating numerous aspects of the experience by producing and analyzing divergent types of data. As such, a set of sub-research questions were specified: **a) What aspects of Virtual Reality should be considered when conducting holistic User-Experience evaluations? b) How well different methods succeed in measuring user-experience of the virtual reality headset? c) How multiple methods can be efficiently utilized within a single Virtual Reality evaluation?**

## USABILITY TEST

The goal of this study was to conduct a holistic user-experience evaluation on a Virtual Reality Headset and acquire results and suggestions that can be used to improve and refine the experience. As evaluating every feature of the experience would be very challenging, the tasks of the experiment revolved around the tutorial included in the evaluated system, interaction with the user-interface, a function that assists the user in exiting from the content and the video player of the headset. To be able to flesh out potential issues that users experienced when interacting with the experience, usability testing and other data collection methods were chosen as a way to obtain information from the participants. In a common usability testing session, tasks are usually introduced to the users one after another, with breaks taken between each task to ask follow-up questions. This approach, however, is unideal for usability testing of Virtual-Reality content, due to:

1. Interrupting the natural flow that participants achieve when interacting with the content of the headset by making them constantly remove and wear the headset.
2. The uneasiness of conducting a "blind" interview, where participants would have to answer questions while wearing the headset and pausing the content.

## System

The head-mounted display that is going to be used herein is Pico G2. Pico G2[12] is a fully immersive, stand alone, 3DOF Virtual Reality system which was manufactured and released by the company Pico Interactive in 2018. As this headset is considered to be stand alone, there are no requirements for connecting it to a computer to operate. The user can put the charged headset on and start using it. Interaction with the Virtual Reality content happens through the head-movements of the user. Pico G2 features a 3K binocular resolution display, 90HZ refresh rate and a total weight of 268g.

The headset included the possibility of watching a set of videos and navigating the categories that the videos were organized into. Upon booting the head-mounted display, users are required to complete a tutorial before interacting with the main content of the headset. The tutorial familiarizes users with some of the functions of the headset, such as how to select objects by moving their heads and how to exit from the content being viewed, while additionally informing users about the locations of physical buttons on the headset which can be used to interact with the experience. When the tutorial is complete, users are able to freely navigate the experience without further interruptions.

In the starting area, three window interfaces appear within the peripheral vision of the user showcasing the available content. Users can then move their heads to browse through categories of videos and games. When a category is selected, the users are able to see the options

[12] The following information was retrieved from the official Pico Interactive website, for more information on the Pico G2 please visit: https://www.pico-interactive.com/g2.

within that category before initiating them. Depending on the type of content, users have the option to adjust volume, navigate in a video being watched, exit from their current choice, or use the physical buttons of the headset to interact with the content.

## Participants

It was decided that all participants to be at least 30 years old. This age group was chosen as people over 30 years old potentially have greater economic flexibility to rent or buy a Virtual Reality headset due to their income. The study had 14 people participating. The number was deemed to be adequate as it has been shown that a total of 5 to 8 participants can expose around 80% of usability problems through testing with users [Nielsen and Landauer 1993]. As usability testing was a large part of my holistic UX evaluation for this specific thesis, 14 participants were enough to produce qualitative and quantitative data for this experiment.

There were 9 male participants and 5 female participants. Their ages ranged from 30 to 51, with a mean age of 34 years. Nine out of fourteen participants reported that they were 'quite interested' in Virtual reality, half of them had previous experiences with virtual reality, and eight participants wore glasses. Nine out of fourteen participants reported spending between four to eight hours a day using different kinds of technology such as computers and smartphones. Lastly, all of the fourteen participants reported having no cognitive or physical impairments that could impede them from taking part in this study. A more detailed description of the demographics is shown in Table 2.

**Participant characteristics:**

| | |
|---|---|
| *Avg. age (in years)* | 34 |
| *Percent male* | 64% |
| *Percent female* | 36% |
| *Quite interested in Virtual-Reality* | 64% |
| *Prior Virtual-Reality experience* | 50% |
| *Percent wearing glasses* | 57% |
| *Using technology between 4 to 8 hours every day* | 64% |

*Table 2. Demographics of participants.*

## Data collection

Data was gathered via questionnaires, the think aloud method, usability tasks, notes made during the session, and interviews. The set of questionnaires consisted of:

1. **A pre-experiment questionnaire**

   Primarily focused on understanding how technologically-savvy the participants were. Additionally, the pre-questionnaire had questions about the general level of interest of participants regarding Virtual Reality, as well as how often they interact with technology in their daily life. There were no questions in the pre-questionnaire that were related to the educational level or profession of participants (see APPENDIX A).

2. **Pre & Post Comfort questionnaire**

   The comfort questionnaire asks the participants about different aspects of their current condition, each having a set of 4 answer options (see APPENDIX B). The Comfort questionnaire was provided by the company and it is one of the tools that they use to evaluate how participants feel before, and after their experience with the headset.

3. **System Usability Scale questionnaire**

   The System Usability Scale was originally created by John Brooke in 1986 [Brooke 2013], with the purpose of offering to the users statements that describe the usability of a product [Brooke 1996]. To answer the questionnaire questions, users have to respond a Likert style scale that is comprised of 5 choices that range from strongly disagree to strongly agree. After participants answer their questions, a SUS score is calculated to evaluate the system. If the resulting SUS score is above 68, then the system is thought to be above average. In the case of this thesis, the SUS questionnaire was used to evaluate satisfaction with the product (see Appendix C).

**4. Presence questionnaire**

The third questionnaire was the Presence Questionnaire. It was created by Bob Witmer and Michael Singer in 1992, with the goal of measuring the amount that individuals experience presence inside a virtual environment and additionally, how specific elements of the experience can possibly influence the intensity of the experience [Witmer and Singer 1998]. To answer the questionnaire, users have to fill in a scale that offers 7 choices with different descriptive labels. As such, presence in the questionnaire is considered to be the averaged sum score from the following subset of categories: “Realism”, “Possibility to act”, “Quality of Interface”, “Possibility to examine” and “Self-evaluation”. The questionnaire questions correspond to the above categories. In order to produce the total sum of the questionnaire each category need to be calculated independently amongst all participants first. For example, to calculate the overall score of “Possibility to act”, you add the sum of the questions 1, 2, 8 and 9 from the questionnaire. This produces the score for the category “Possibility to act” for one participant. Similarly, the rest of the sums are calculated for all categories. The total sum of the averages of all categories is the final Presence score (see Appendix D).

This study used a combination of methods that collected both quantitative data from result-based activities, and qualitative data through direct observation and personal notes. The researcher recorded both qualitative and quantitative measures in the form of notes in a separate checklist (see Appendix G) during the testing session. Table 3 below presents the quantitative and qualitative measures that were recorded are explained in more detail.

| **Quantitative measures**: | **Qualitative measures:** |
|---|---|
| Whether assistance was provided to the participant. | Think-aloud comments made by participants. |
| Time it took to complete a task. | Personal notes of the moderator in form of major problems participants had. |

| | |
|---|---|
| Whether a task was considered to be successful or not. | Responses to follow-up interview questions (see Appendix E). |

*Table 3. Quantitative and qualitative measures recorded from the testing session.*

## Procedure

### Pilot test

Before the actual experiment, a pilot test was conducted with one participant and several small changes were made to the experiment. The order of the usability tasks was rearranged to avoid introducing participants to certain features of the headset before a certain task. Additionally, spelling corrections were made on the questionnaires and interview questions were rephrased to avoid bias in the collected feedback.

### Preparations for the test

Before the experiment began, the researcher had prepared in advance the equipment that would be used. As such, the researcher had to:

1. Set up the VR headset to connect via Wi-Fi and Miracast to an external laptop.
2. Connect the external laptop and the headset together via Wi-Fi.
3. Activate the audio & video recording equipment after participants had signed the consent form.

Screen sharing was applied to display the in-headset actions of participants to an external laptop, which allowed the researcher to record what the user was doing whilst interacting with the headset. Additionally, a mobile phone was used as a backup to capture the audio from the testing sessions. To avoid participants getting distracted from the Virtual-Reality experience the experiment took place in sound isolated rooms in the library of Tampere University. The researcher kept notes in the form of an observation checklist (see Appendix G) recording both qualitative and quantitative data from the session. The checklist included information of successful tasks and their duration, prompts made by the moderator, major problems encountered by participants and think-aloud comments.

**Testing session**

Before the beginning of the experiment participants had to first sign a consent form (see Appendix E). Next, they were asked to fill in the pre-experiment questionnaire regarding their enthusiasm and previous experiences in Virtual-Reality. The researcher then remained with the participant in the testing room throughout the session and introduced the tasks. The tasks were:

1. *"I would like you to adjust the headset so that the image feels comfortable".*
2. *Have users watch the tutorial.*
3. *Select Culture and Sports by using the selection button of the headset.*
4. *Explore the headset on your own.*
5. *Go to category Cinema & TV. Select Movies. Return to the homepage.*
6. *Go to category Travel & Relaxation. Select Relaxation. Select "Echoes in White".*
7. *Go to category Cinema & TV. Select Movies. Select "What do you desire?*
8. *Go to Culture & Sports. Select Music. Select "1969". Go to music. Return to homepage.*
9. *Go to Category Shopping. Select Electronics. Select Power Pack.*
10. *Center your vision by pressing the secondary button. Recenter your vision three more times.*
11. *Go to Category Culture & Sports . Go to Music. Select the video "Me and my Drummer". Click play.*
12. *Go to Category Travel & Relaxation. Select Dream Beach. Click Play.*
13. *Go to Category Kid's Corner. Select "Secrets of Gravity". Click Play.*
14. *Without touching the volume button on your headset try to increase the volume of the video to 10.*
15. *Go to category Travel & Relaxation. Select Diving. Select and play "Phillipines underwater". Return to the homepage.*
16. *Go to Category Cinema & TV. Select Movies. Select and play "Tadeo Jones" Skip to minute 3:00 of the video without pressing the selection Button.*
17. *Start a video of your own choosing and pause it. Skip to minute 0:30 of the video. Then skip to 1:45.*
18. *Center your vision by pressing the secondary button.*
19. *Explore the headset on your own.*

20. *(If participants did not select environment in the previous task). "You are currently a passenger on an airplane in your headset. Try to change your environment to a beach".*
21. *Try to re-initiate the tutorial of the headset.*

The tasks were presented in a fixed order, one task after another, and the participants were observed by the researcher while performing tasks. The reason for this is because the researcher did not want to introduce the participants to certain functions of the experience before the participants interacted with other aspects of the experience due to them being connected in a specific way. The moderator additionally encouraged the participants to think aloud as they interacted and navigated through the virtual environment of the headset, which gave the opportunity to the researcher to identify potential pain points that participants had. Once the participants had finished all the tasks, they were given the Post-Comfort questionnaire, the System Usability Scale questionnaire and Presence Questionnaire to fill in that order. After all the questionnaires were completed, the participants answered a number of questions in a semi-structured interview about their experience in interacting with the headset, as well as their overall impression of the experience. This semi-structured interview (see Appendix E) aimed to fill in the remaining gaps of data.

## 4) RESULTS

This section is divided into two parts presenting the respective results: a) Usability testing, b) Questionnaires.

### Usability Testing

Results from usability testing are presented in table 4 below. Most of the participants were capable to successfully complete their tasks without the aid of the moderator. However, tasks 3, 7, 8, 9 10, 20 proved to be the hardest to complete, with numerous participants being unable to fulfill the task. Even though all 14 participants were able to complete the remaining 15 tasks of the experiment, a diverse range of usability problems were encountered throughout the think-aloud and the follow-up interview session.

| *Tasks* | Successful attempts (n=14) | Failed attempt(s) | Average successful completion time | Participants that required assistance |
|---|---|---|---|---|
| *Task 1* | 14 | - | 52s | - |
| *Task 2* | 13 | 1 | 1m 19s | 1 |
| *Task 3* | 10 | 4 | 29s | 2 |
| *Task 4* | 14 | - | 4m 5s | - |
| *Task 5* | 14 | - | 26s | - |
| *Task 6* | 14 | - | 23s | - |
| *Task 7* | 5 | 9 | 27s | - |
| *Task 8* | 5 | 9 | 29s | 7 |
| *Task 9* | 12 | 2 | 24s | - |
| *Task 10* | 10 | 4 | 22s | 3 |
| *Task 11* | 14 | - | 1m 48s | - |
| *Task 12* | 14 | - | 1m 47s | - |
| *Task 13* | 13 | 1 | 45s | - |
| *Task 14* | 13 | 1 | 40s | - |
| *Task 15* | 14 | - | 1m 26s | - |
| *Task 16* | 14 | - | 58s | - |
| *Task 17* | 14 | - | 1m 35s | - |
| *Task 18* | 14 | - | 10s | - |
| *Task 19* | 13 | 1 | 1m 30s | - |

| | | | | |
|---|---|---|---|---|
| *Task 20* | 8 | 6 | 43 | 6 |
| *Task 21* | 14 | - | 40s | - |

*Table 4. Results from the usability testing tasks.*

## Usability problems

The most important areas that the participants encountered usability issues are summarized below in four categories: tutorial, user-interface, exit-function and video-players. Each category is broken down into sub-sections describing negative and positive aspects according to the comments made by participants.

### TUTORIAL

*How well the tutorial assists participants with understanding the functions of the headset?*

Participants reported having issues with reading the slides of the tutorial related to the functions of the headset: “The info on the second slide went through before I could read it, I started reading the text on the right side but I had no time to finish reading,” reported one participant. Similarly, another participant commented on how they were unsure whether the slides were proceeding through their actions or not: “Was I supposed to look at the text and wait for the next slide? Or should I push a button to go forward?

Other participants noted that the tutorial explained functions of the headset with unfamiliar terminology: “What does it mean that it re-centers your vision?” mentioned one participant, similarly another participant exclaimed how “the name of the round button was not familiar to me at all”.

Lastly, as shown in Task 3, numerous participants ended up confusing the two selection buttons of the headset and required help from the moderator in order to interact with them. Some of them reported being unsure about the location of the buttons: “Is the concave button up or down?,” while other participants could not find the buttons at all: “I don’t know if I found it, I don’t know which one it is”, or they could not remember the name of the buttons: “Is this the one the oval button? I could not remember the name of the button”.

*How clearly is the information of the tutorial presented to the users?*

Participants reported struggling with understanding from the tutorial which buttons they had to use for specific functions. This confusion was especially noticeable in task 10, where participants required help from the moderator to understand where the buttons were in order to use them. “When I saw the image of the tutorial it was difficult to understand which button was the right one," said one participant. Another noted that “I did not think that I had to keep pressing the concave button, I did not remember that from the first tutorial”. One participant mentioned how the tutorial should familiarize them with both buttons, by observing that “for the oval button they could make you choose some things also in order to get familiarized with it before you start”.

*The tutorial as an entry point to the headset experience*

“Well nobody really wants to look at tutorials, but it was short enough, and it’s good to know how this thing works,” commented one participant. Another suggested that “in the beginning, when a person uses something for the first time, their concentration is not there 100%”. Moreover, other participants commented that while the tutorial explains some functionalities of the headset, it disregards explaining others: “Should the function of selecting videos be in the tutorial maybe?” remarked one participant. Similarly, another one explained that the tutorial “was focused more on the hardware of the headset and then you are thrown into the environment without telling what's more there”.

*Aesthetic look of the tutorial*

“It was nice, sharp, distinct, and the fonts really helped with understanding were things are especially with the buttons,” noted two participants during the interviews. Similarly, several other participants mentioned that the Virtual Reality experience was new and impressive. On another note, some participants considered the overall look of the tutorial to not be so engaging. “Old-school, it seemed that they just wanted to get the information through and it could have easily been an image on a computer with a tutorial,” commented one participant. Another participant expected the tutorial to be more “real” as she explained that “when you are in VR you want to look at things and say that this is real, but it was a plain kind of like slide.” Another commented that “it looks incomplete. The instructions were okay and straightforward, but the presentation was like a demo version.”

## USER-INTERFACE

### *Choosing content by looking at the selection icon*

Participants were instructed to explore the environment on their own in task 2, as well as access various videos in the tasks 4 to 9. Observing how participants interacted with these menus exposed multiple issues. The icon that allowed them to select respective videos was the first complication. For example, numerous participants explained that they had trouble with understanding the purpose of the icon that was used to select the content: “I also did not think that I could select things with the eye before I tried**,”** commented one participant. “When I wanted to select that movie I did not look at the icon, I thought that just looking at the square would be enough,” noted another one. Similarly, other participants pointed out that the small size of the icon did not assist with its overall detection and thought that “the icon was part of the overall picture”. On a similar note, content selection sometimes happened accidentally as some participants glanced over the icon while using the headset: “It was accidental that I chose this first music video, after that I understood that that is the way I chose the video”. However, it was observed that it became easier for participants to understand how the icon works as they got more familiarized with the headset.

### *Choosing content via the breadcrumb option*

Participants did not seem to observe and use the breadcrumb option as much as the back-arrow function of the experience in tasks 7 and 8. Specifically, in task 7, several participants continued selecting content via gazing at the selection icon, whereas in task 8, the moderator had to directly prompt multiple participants to look for a different way of selecting the content. One participant remarked that “I could not figure out the breadcrumbs, I was in that category and I did not know that I could highlight the breadcrumbs and it would take me to music.” Other participants commented that “the small size of the breadcrumbs made them difficult to spot” and that “the selection area around the breadcrumbs was too small when they looked there”. Another participant gazed at the breadcrumbs too quickly and their selection did not occur fast enough: “I looked at the breadcrumbs and it turned white, but my choice did not go through though”.

*Discoverability of the environments section*

The menu that the participants first encounter after completing the tutorial includes a section that allows users to transform surrounding environment to a different type of environment. However, not all 14 participants noticed this area of the interface at the beginning of the experiment, even when they were left to explore on their own during task 4. Six of them failed to discover it during Task 20, and thus the moderator had to request for them to search for the option that allows them to change their surroundings. On that note, a participant remarked that "the environments on the right side do not seem to be selectable."

## EXIT-FUNCTION

*Detecting the exit-function during video playing*

In addition to navigating menus and selecting videos, the participants were also required to navigate by exiting from videos. This brought its own challenges. While almost all of the participants were able to discover the exit function without inconveniences or help from the moderator in tasks 11 to 14, several of them remarked that they found the exit-function difficult to detect. "I did not see the cube to return back home, it was not noticeable, perhaps because of the color," said one, "it was just there all the time but transparent," commented another during the interview. Others noted that they could not find or see the exit-function all-together: "I could not see it at all when I was looking at the butcher video, I did not see it." Another participant noted that if she "would not have known that the home cube was always in the same location it would probably have been hard," when explaining her attempts to exit from videos during the interview.

However, several participants explained that they needed some time to familiarize themselves with the exit-function of the interface, saying "Now I am used to it, now I do not see this as a problem", and "once you get the idea with the cube and the button that from there you have your functions you pick the logic very easily," and "I knew it was there I just had to look closer." One participant commented that "the exit function is always at the same place." Another remarked, "it was easy because it was always in the same direction, left up place in my vision."

*Actions on the exit-function*

Some participants reported having trouble recognizing a certain icon on the exit-function interface of the headset. “If there would not have been the task to center the vision, I would not have used the icon, or would not have connected it to the task at hand," exclaimed one. Additionally, another participant remarked that they were confused by the icon and “did not understand what the icon was”.

## VIDEO PLAYERS

### *Interacting with the video player bar*

A few participants noticed that the video bar did not appear in consistent locations from video to video. “I think in one video I had to search for the video bar since I could not find where it was”, and “in other videos it just appeared there (pointing at a certain point).” Similarly, one participant remarked that she felt that the video bar was higher than on the rest of the videos and also covered the action that was taking place: “the video player bar comes too high up, it should be lower because now it is kind of bothering me because I want to see where the turtle is going, and the bar is appearing.” One participant reported experiencing strong neck pain and proceeded to show to the moderator the extent of which he had to strain his neck in order to interact with the video: “it is way too low and I have to stretch my neck like this which makes it really uncomfortable for my neck”.

Additionally, several participants reported encountering problems in form of a glitch when trying to grab the round markers that control video and audio and change their positions in tasks 14, 16 and 17. “You see? I am looking at the video player bar and the video is not responding, it is not responding now again,” commented one participant. One participant noted in the interviews that she “had to approach the video bar from below or from above to control the video bar”. Similarly, another explained that, “When I put the cube further down the line, it didn’t move,” and “I had to look away and then look back at the line which was annoying.” By observing the video recordings of the participants and then replicating the problems it was shown that the bar successfully registered the selection of the participants when they gazed away from the bar and re-approached it from a different direction, as moving their gaze horizontally on the bar and trying to change the marker location did not register their choice.

### *Feedback information on the video player bar*

Several participants remarked that they felt that trying to adjust the marker in the videos in tasks 16 and 17 was one of the most challenging tasks they had faced due to the lack of feedback. “I would move my head on the video time-bar and the video would load, but I would not know which minute I was going too,” said one. Others noted that the video bar “should show the number of the video somehow,” and noted that “either you put time indicators on the video, or you show by numbers where you are tracking to”. Similarly, some participants drew examples from other platforms such as Netflix or Youtube, to highlight their familiarity to the time feedback feature. “In Netflix when you look at the bar and it shows you the hours and minutes,” exclaimed one participant. Similarly, participants explained that it was unfamiliar for them to find the option that allows you to control the audio in the exit-function interface. One participant said that “In Youtube you have the sound icon next to the play button,” and another said that “I think that the sound is unnatural to be on the exit-function menu”.

## OTHER OBSERVATIONS

*More loading animations as a form of feedback*

Participants during the interview noticed that the loading animations appeared for only some of the features of the headset. “The big buttons on the left have the blue selection progress bar but the selection symbol does not,” said one participant. This absence of feedback lead to confusion during the experiment for one participant who “thought that maybe the video content pictures would have a bar or some other sign that shows me that something is happening,” and “the same with the play and selection icon button, it would be nice to have a loading bar or some other indication.” Others mentioned that other features of the headset should include loading animation as feedback as well: “You could add the loading bar animations to the breadcrumb to inform the people about loading,” exclaimed one participant.

*Graphics of the headset*

Half of the participants expressed that they did not consider the graphics of the headset to be high in quality during the interviews. “Images could be sharper as the pixels are visible,” said one participant. Similarly, another participant explained that “the pixels could be clearer”. Several others explained additionally that they expected in general the graphics to be higher in quality. “I really expected much more," observed one participant. Another noted that he thought that "still looks the definition is too low" and "maybe I am expecting too much". However, these impressions could be the result of the low quality of the video in task 15, "Philippines underwater" which ended up biasing the opinion of participants. "In this video (Philippine's underwater) the quality is a bit poorer than others," noted one participant. Another one remarked that "the video is too blurry," and "the lines seem to be shaky". Moreover, other participants compared the Philippine's underwater video to the "Dream beach" video during the interviews were they referred to the Dream beach "as nicer, and more realistic". On another note, other participants believed the "graphics are sharp compared to other headsets" and that "the videos could be of higher quality but are still really impressive."

*Immersion*

Close to half of the participants felt that the experience looked realistic which helped them get immersed in the experience. "It made you believe that you were there" and "it made me forget my real environment," described one participant. Another noted that she "really felt that she was in a real space". A third participant explained that "the ability to look around" was something that he always wanted to have.

Moreover, two participants described that becoming immersed in their experience assisted in forgetting about their worries and physical discomfort that they were feeling. "The experience diverted my attention and the thoughts and stress that I had in my brain," commented one of the participants during the interview. The other participant reported to his surprise that the migraine he had before the experiment had disappeared after the experiment. On another note, other participants explained how the experience did not feel so real due to the "resolution" and "quality of the image," whilst adding that "in an experiment, you don't have the flexibility to immerse yourself more" due to the short amount of time and comments made by the moderator.

Additionally, participants commented that they did not find engaging watching 2D videos in a virtual reality environment setting. One participant commented that in the movie theater experience "the video content was like a 2D video but in a virtual reality environment". Another explained that he was "expecting to be inside the movie in 3D". Participants contrasted 2D videos to the dream beach and the dome cinema where the content felt "different than looking at a flat screen" due to the opportunity to "see something everywhere, as it was more than 180 degrees.".

Furthermore, participants during the experiment commented on the absence of a digital body in the experience. "Oh my god I have no body," remarked one participant, or "I keep looking for my hand and it's not there hihi," said another. On a similar note, participants commented on the absence of other "people" around them as the experience was simulating a space were participants expected to see others. "Why am I alone?" and I would like to have people here," remarked one participant. Another commented that "it is a little bit scary because the surrounding environment is empty."

*Physical discomfort*

"The cushion started to become a bit warm and sweaty during this test" and "the lenses were a bit foggy" explained one participant at the end the experiment. A handful of others pointed out the weight of the headset after usage. "The headset was a bit heavy on the cheekbones," said one participant. Another said that "I can feel that my neck muscles have worked" and "I am not sure if it would make my neck hurt more if I used it longer". Other participants expressed that the headset "did not fit their head properly" or "it felt a little bit of pressure on my face because of the strap". At this point it should be noted that the moderator did not notice the possibility of adjusting the tightness of the strap of the headset until very late in the interviews, which potentially impacted the experience of comfort of users.

## Questionnaires

This section presents the results of the questionnaires that were used in the experiment: a) System Usability Scale, b) Comfort questionnaire and c) Presence questionnaire.

### System Usability Scale (SUS)

The overall System Usability Scale (SUS) grade for the headset was 73, relatively higher than the standard score of 68. The average SUS score for the participants that had used Virtual Reality before (n=7) was 76. The average SUS score for participants without any previous Virtual Reality experience before (n=7) was 71. The raw data of the SUS questionnaire can be found in Appendix I.

### Comfort questionnaire

The comfort questionnaire was administered to all 14 participants before and after the experiment. It was observed that out of all 16 categories, 9 seemed to showcase evident changes in how participants were feeling. Furthermore, the raw data from the comfort questionnaire and the results before and after the experiment can be found in Appendix H. The 9 categories that presented evident changes in symptoms to participants are briefly explained below.

**GENERAL DISCOMFORT**

The **overall general discomfort** seemed to have increased for the majority of participants after the experiment.

| | BEFORE | AFTER |
|---|---|---|
| NONE | 12 (85%) | 7 (50%) |
| SLIGHT | 2 (15%) | 5 (35%) |
| MODERATE | 0 | 2 (15%) |
| SEVERE | 0 | 0 |

**EYE STRAIN**

Similarly, it was observed that a considerable number of participants experienced symptoms of **eye strain** after the experiment was over.

| | BEFORE | AFTER |
|---|---|---|
| NONE | 11 (79%) | 6 (43%) |
| SLIGHT | 3 (21%) | 4 (29%) |
| MODERATE | 0 | 3 (21%) |
| SEVERE | 0 | 1 (7%) |

This was expressed by participants during the experiment as certain logos and menus during videos the experience seemed to appear close to their eyes creating eye strain. “The text that appeared from centering vision appeared very close to my eyes and felt uncomfortable,” explained one participant. “When the logo of the company appears, as if it is right in front of my eyes, and not the way you see the rest of the videos,” remarked another during the video “Blood-Moon Child”. Other participants during their second viewing of the tutorial section preferred the differences in distance, size and contrast between logos and menus appearing in the videos and the tutorial. “This is a better distance to read text, because it was a little bit further from me it was easy to look and understand,” explained one participant. “This logo (the company logo appearing in the tutorial) is better, it is bigger and is further and the background is not black, so the contrast is not so big,” remarked another.

**NAUSEA**

Almost half of the participants started experiencing symptoms of slight **nausea** after the experiment was over.

| | BEFORE | AFTER |
|---|---|---|
| NONE | 14 (100%) | 9 (65%) |
| SLIGHT | 0 | 5 (35%) |
| MODERATE | 0 | 0 |
| SEVERE | 0 | 0 |

**DIFFICULTY CONCENTRATING**

On another note, it was observed that participants after the experiment were able to **concentrate better** than they could before the experiment started. Similarly, two participants reported that their feelings of stress and migraine had disappeared in the interviews.

| | BEFORE | AFTER |
|---|---|---|
| NONE | 9 (65%) | 10 (73%) |

| | | |
|---|---|---|
| SLIGHT | 4 (29%) | 3 (27%) |
| MODERATE | 1 (6%) | 0 |
| SEVERE | 0 | 0 |

**BLURRED VISION**

Participants reported that there was an overall **increase of blurred vision** of the participants after the session.

| | BEFORE | AFTER |
|---|---|---|
| NONE | 12 (86%) | 9 (65%) |
| SLIGHT | 2 (14%) | 4 (29%) |
| MODERATE | 0 | 1 (6%) |
| SEVERE | 0 | 0 |

**DIZZINESS WITH OPEN EYES**

Similarly, **dizziness both with open and closed eyes** after the experiment had increased.

| | BEFORE | AFTER |
|---|---|---|
| NONE | 13 (93%) | 9 (65%) |
| SLIGHT | 1 (7%) | 4 (28%) |
| MODERATE | 0 | 1 (7%) |
| SEVERE | 0 | 0 |

**DIZZINESS WITH CLOSED EYES**

| | BEFORE | AFTER |
|---|---|---|
| NONE | 14 (100%) | 10 (72%) |
| SLIGHT | 0 | 3 (21%) |
| MODERATE | 0 | 1 (7%) |
| SEVERE | 0 | 0 |

**VERTIGO**

Lastly, there was an overall **slight increase in the feeling of vertigo** in participants after the experiment was over.

| | BEFORE | AFTER |
|---|---|---|
| NONE | 14 (100%) | 11 (79%) |
| SLIGHT | 0 | 3 (21%) |
| MODERATE | 0 | 0 |
| SEVERE | 0 | 0 |

***Presence questionnaire***

The average total score of Presence of the 14 participants was 79. Similarly, the average scores of the subset categories that contribute to Presence and their average scores are presented in Table 5 below.

| | Average |
|---|---|
| Total Presence score: | 79 |
| «Realism» | 24 |
| «Possibility to act» | 22 |
| «Quality of interface» | 16 |
| «Possibility to examine» | 5 |
| «Self-evaluation of performance» | 12 |

*Table 5. Results of the Presence questionnaire.*

# 5) DISCUSSION

This chapter will answer the research questions of this study by presenting the aspects of this experience and showcasing how well the evaluation methods succeeded in providing data. Furthermore, the aspects and methods discussed are summarized in a table followed by a list of contributions of this study.

## Research questions

User-experience evaluation in Virtual reality is not yet a fully understood process. Since the technology is still at an early stage of its development developers and researchers have to consider technical aspects (e.g. processing power of computers), affective aspects (e.g. excitement, fun, frustration) and physical ergonomics (e.g. comfortable to use HMD's or controllers) of the technology that can influence the experience of users. These challenges make the scope of evaluating user-experience in Virtual Reality very broad.

The research objective of this thesis was to **investigate how we can perform a holistic user-experience evaluation on a Virtual Reality headset**. The term *holistic* refers to something that covers all aspects of a certain topic. In Virtual Reality we can consider numerous aspects of an experience, including physical ergonomics, interaction and content. In this study however I do not give a theoretical definition of what holistic user experience is in the context of Virtual Reality. Rather, I describe aspects of Virtual Reality that affect the user-experience in this particular evaluation, as there are endless details that one could be considered as part of a holistic view of user experience in general. In this particular experience they included interaction with the user interface, technological characteristics of the system, presence and ergonomics which will be described in this discussion. Moreover, as Fenton and Pfeeger explain: "you cannot control what you cannot measure and you cannot measure what you cannot define" [1997, p.41]. This led to the creation of the *first sub-question* of this thesis: **What aspects of Virtual Reality should be considered when conducting a holistic User-Experience evaluation?** Secondly, as an evaluation was conducted as part of this thesis, it was deemed necessary to examine **how well different methods succeeded in providing data to measure the holistic user-experience of the virtual reality headset, and how these methods can be efficiently utilized within one Virtual Reality evaluation.** The following sub-chapters will attempt to answer the research questions mentioned above.

## Aspects of the experience

The first sub-question acted as an incentive to understand the holistic aspects that affect the user-experience of this evaluation. This involved characteristics of usability, user-experience, virtual reality and understanding how they influence each other. Furthermore, to answer this research question, this section combines the collected information from the chapter User-Experience and Virtual Reality as well as observations from the experiment of this study.

As Forlizzi and Battarbee [2004] explain, User-experience has been correlated to both pragmatic and hedonic aspects of a technology being used. The field of Human-Computer Interaction tried to reduce the increasing complexity of technology by creating systems that allow the user to achieve their goals with ease, while keeping a steady focus on efficiency, effectiveness and satisfaction of using a product. This is also known as usability. Zimmermann [2008] explains that usability is a popular and well-established term for measuring the quality of products. However, the author describes how over recent years, new approaches such as pleasure, emotional usability, and hedonic aspects of technology were subsumed alongside usability in what is know known as the field of User-Experience.

Even though academic research regarding User-experience in Virtual Reality is sparse, usability has been the term used in the Virtual Reality literature and is the key that connects User-experience and Virtual Reality together. Examples such as the MAUVE system explain usability in Virtual Reality as a form of criteria that developers and researchers can consider when creating Virtual Reality applications [Stanney *et al.* 2003]. In this study, the following criteria mentioned in MAUVE were identified during the testing and analysis of the data:

1. Wayfinding
2. Object selection and manipulation of the Virtual Reality content
3. Technological characteristics of the system
4. Comfort
5. Simulator sickness

In regard to wayfinding, Furnas [1997] describes that the user should be able to strategize the nearest path to reach their objective by using their vision. In this study, the participants had time to observe the user interface and then discover how to do the selections required in the tasks. Similarly, Wickens and Baker [1995] state that the virtual environment should bear enough clues to allow the user to know where they can go and how they can achieve that. In the experiment, the user interface provided cues that could guide the user in form of selection icons and breadcrumbs. However, some participants were not able to understand the functions required to interact with the content. This is described in MAUVE under the category of object selection and manipulation. Kalawsky [1999] explains that the affordance of selecting content in Virtual Reality should always be apparent and unambiguous. During the interviews and the testing session participants described that the way the selection function looked did not help them recognize its purpose. Additionally, participants had trouble distinguishing and detecting the exit function in certain videos of the experience due to its transparent color and the contrast of the video.

Furthermore, Ware and Balakrishnan [1994] state that the system should have small selection latency and fast response rates to ensure a smooth and natural object selection throughout the experience. However, observations from the study contradicted this statement to some extent. Some participants were observed to accidentally select content while looking around the user interface when trying to complete tasks. On the other hand, one participant was not able to understand the function of the breadcrumb when he gazed at it because the system did not respond quickly enough. Other participants in the interviews praised the responsiveness of the headset. As such, there needs to be a certain balance between fast response rate, dwell time and selection latency depending on the Virtual Reality system at hand.

Other criteria in MAUVE describe technical characteristics of the systems that developers should keep in mind including stereoscopic display support, spatial resolution, field-of-view, update rates, refresh rates, and user comfort and acceptance. Kalawsky [1999] explains that the users can easily detect visual inconsistencies that can influence the experience. In the study, participants reported blurred and shaky lines in parts of the user interface and videos. Additionally, participants commented on videos that they thought to be of low quality, and in the interviews compared them to ones that were considered to be of better quality. The video “Philippines underwater” was brought up by some of the participants when they described the impression of graphics of the experience. The low-resolution quality of that video made it difficult to understand what was happening in the scene due to the video being very pixelated in specific parts. Consecutively, the visual output of the system can influence the aspect of presence and immersion aspect of the experience.

Another crucial criteria in MAUVE is comfort. Developers and researchers should be aware of the physical risks potentially involved in the experience of the user, as health and safety can directly impact the experience of users. In this study participants described that they felt pressure from the straps of the headset after some time. Others experienced eye strain due to certain logos and menus appearing too close during the testing session. Some participants described that the constant head-movement that was required to select content became tiresome after a while, and one participant experienced severe neck straining when trying to interact with the video player. Factors that cause discomfort, such as continuous usage and weight of the headset, can impact how the users feel during a usability testing session as well as during normal use. Durlach and Mavor [1996] state that users should feel comfortable when wearing a head-mounted display to avoid pain and unpleasantness that can create negative thoughts and emotions. Another criterion that can influence comfort is simulator-sickness. In this study, the comfort questionnaire was able to show the differences in symptoms related simulator-sickness before and after the testing session (see APPENDIX H and comfort questionnaire results).

As the characteristics that can potentially influence evaluations in virtual reality are numerous, the analogy of the puzzle mentioned previously (see end of sub-chapter "User-Experience and Usability in Virtual Reality", in chapter 2) can help to put ourselves in perspective. In the puzzle analogy, user-experience is the overall puzzle. The usability characteristics that have been discussed in this thesis, such as interaction techniques, presence, comfort, simulator sickness and ergonomics are jigsaw pieces that help us complete the puzzle. Depending on the area of the puzzle that we are currently interested in evaluating (e.g. the aspect of the Virtual Reality experience that we want to investigate), we need to identify the puzzle pieces that can help provide us with information to achieve our goal (e.g. the usability characteristics that correspond to that aspect of the experience). It is amongst these lines that the answer to this research question lies: **we should identify the aspects of Virtual Reality that are relevant to our individual case when it comes to conducting holistic user-experience evaluations.** This is important as the aspects that can influence the experience of a user in Virtual Reality are numerous and it can be difficult to consider or know all of them. In this study only a specific number of aspects could be covered and were considered in the holistic evaluation of the experience. Naturally, in other Virtual Reality applications, different themes of the experience can be more prominent depending on the context and the technology being used. We should always understand the components of the experience that matters to each individual case, categorize them, and study them before proceeding with evaluating them.

### Did different methods succeed in providing data to measure the holistic user-experience and were these methods efficiently utilized in this evaluation?

Organizing and analyzing data provided an abundance of information. Usability testing was able to expose pain points the participants had, but there were technical issues hindering the tests. The questionnaires revealed information, but they were unable to pinpoint issues on a deeper level. These are explained in more detail in the following sections alongside with comments on how well the methods were used.

## Organizing and analyzing the data

Qualitative and quantitative data were collected in this experiment through questionnaires, interviewing, think-aloud comments and personal notes made by the moderator during the experiment. The moderator used a checklist (see APPENDIX G) to write the information during the testing sessions on his laptop. The think-aloud comments made by the participants and observations of the moderator were recorded next to the respective tasks on the checklist as the participants used the headset. After each session, the moderator spent time to write the observations in the checklist. This proved to be very useful and time saving to this research as it assisted with remembering and finding the pain points of users when the data was analyzed and video recordings were examined. Additionally, the interviews were transcribed. After all 14 testing sessions were completed, the researcher started organizing the available data.

The researcher created a matrix table (see APPENDIX H) containing the aspects of the headset experience to be analyzed on one column (such as the tutorial and exit function), and where to get the possible information from the collected data in another one (such as the relevant tasks on the checklist and interview questions). This matrix proved to be useful in this study as it provided guidance on how data should be organized before the analysis started.

Afterwards, similarly to the TBCA method created by Neale and Nichols [2001], data was clustered into categories of similar themes to expose pain points. However, in their research, the collected information derived from questionnaires, whereas in this study the collected data was from interviews, think-aloud comments and personal notes. Additionally, compared to their study, there was only one moderator instead of multiple ones in the testing sessions. Having an additional researcher to assist with the categorization of the data would have made the clustering process faster and more efficient. In this study, the process of organizing the data was the longest and hardest.

One potential reason was that the researcher tried to compile thematically similar data into groups before examining all of the data. This made it difficult later on to change existing clusters of data into new clusters or add new data to the existing clusters. A better approach would have been to start grouping the information after all the data was put in the same place. That way it would have been simpler to go through the information and start creating groups instead of trying to change the existing ones. Furthermore, having another researcher would not only provide help with the clustering process but offer a second opinion regarding the categories, as data is sometimes ambiguous. Discussing and sharing opinions with another researcher would provide an additional layer of objectivity to the grouping and interpretation of the data. After all the raw data was grouped into thematic categories, titles describing the problems were assigned to the groups. This made it easier to identify the pain points and also have the information of the users describing a particular issue in the same place. The researcher then put the titles back into the matrix table in order to know which group provides information regarding a specific aspect of the headset alongside other relevant information.

Figure 3 summarizes the organization and analysis process described in this sub-chapter.

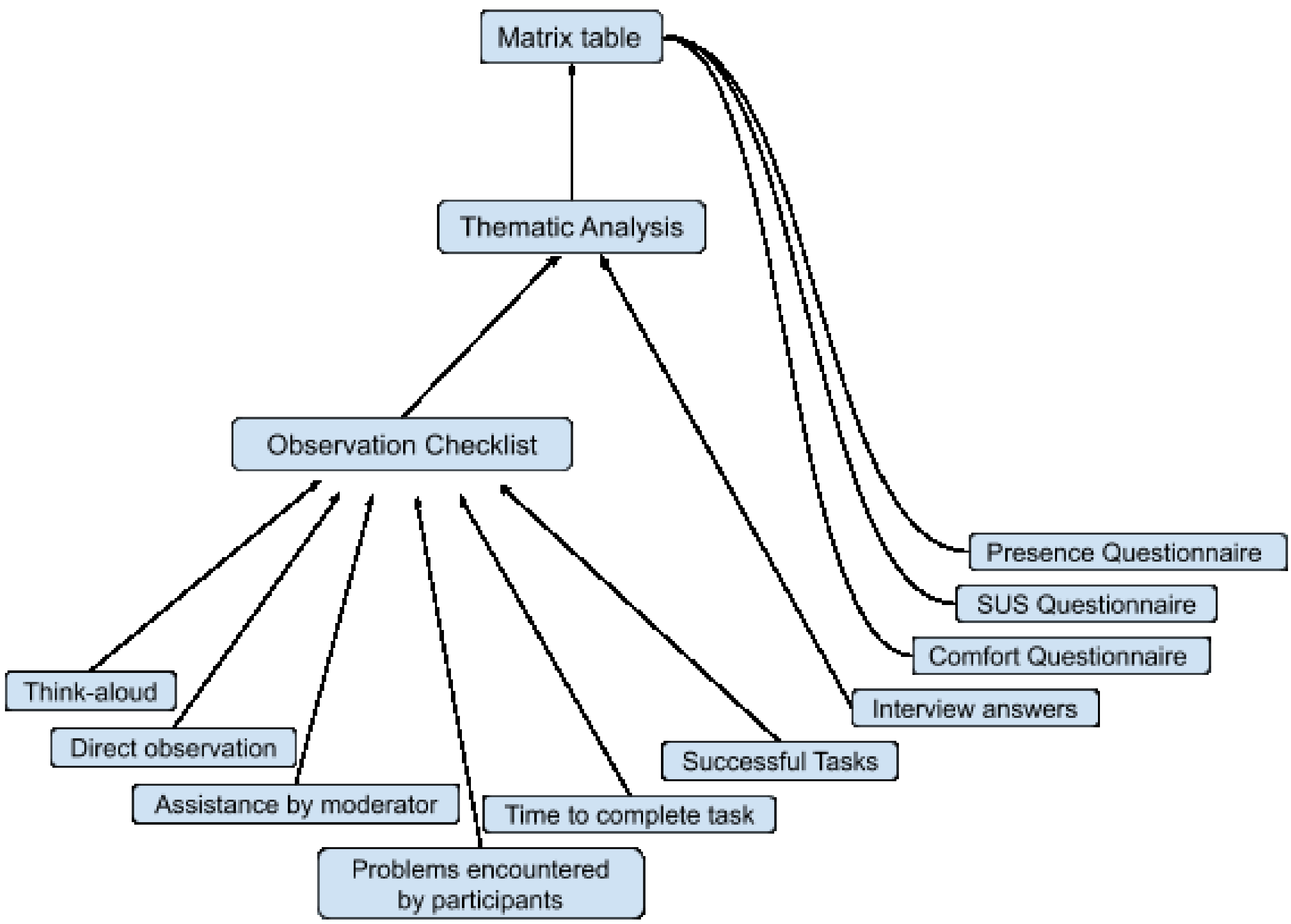


*Figure 3. Organizing and analyzing the collected data.*

## Usability testing

For this thesis a list of available usability evaluation methods was provided (see Table 1 in chapter "Measuring usability" for the list of available usability evaluation methods) The experiment followed the form of a Formative evaluation for Virtual Reality [Hix *et al.* 1999]. I observed that depending on the Virtual Reality system one of the most challenging aspects of conducting a usability test is the inability of the moderator to directly observe what the user is seeing. This problem is especially prominent in Virtual Reality that includes head-mounted displays.

In usability tests a moderator is sitting next to the participant to directly observe the interaction of the user with the system. That way the moderator is able to identify pain points and prompt the participant if needed throughout the testing session. For usability testing with head-mounted displays a possible solution is to project what the users are seeing on a separate laptop which allows the moderator to see what the participant is experiencing in the virtual environment. However, this solution does not come without its challenges.

The head-mounted display could connect to a windows computer through Wi-Fi and Miracast so that it could screen share what the participant was seeing. However, more than once during the testing sessions the devices were unable to connect using an available network and mobile phone had to be used instead. There were also instances where the computer and the headset were connected but no video was projected. Repeated restarts of the computer were required. In general, headsets that require such wireless connections to screen share can be problematic in tightly scheduled testing sessions because of such connection issues. It puts the moderator in a stressful condition where they interact with a waiting participant, while trying to solve the technical problems. This stress can have a negative impact on the focus of the moderator. For example, twice the moderator noticed halfway through a task that he had forgotten to activate the stopwatch.

Another challenge regarding the devices of the experiment was that the screen shared image on the laptop was not shown in full-screen mode. The shared video appeared as a window that covered a small portion of the screen without the possibility to change its size. This made it difficult for the moderator to monitor the shared image to identify issues that participants encountered. The moderator would then have to either rely on his notes, or rewatch the video recording of the testing session to understand the root of the problem.

Another challenge when it comes to usability testing with head-mounted displays is that the headset makes it impossible for the moderator to observe the face of the participant. This limits the ability of the moderator to receive cues from the facial expressions of participants as to whether they are excited or struggling with the experience. In a different context, the moderator could have interpreted the facial expression as notes (e.g. "participants were visibly excited when X happened") or rely on the participants' facial reactions to subtly prompt the participant to investigate a potential pain point during the testing session (e.g. "Is something from the experience troubling you?"). This is especially useful for participants that are not very talkative on their own. Alternatively, the moderator can use video recording equipment to capture the facial expressions or body posture of the participants and examine them later. However, in Virtual Reality usability evaluations the moderator can only observe the body posture of participants or hear what the participants are describing. This was often the case in this study when participants would often make sudden movements. In that moment, the moderator tried to make a note of what happened and ask the participant. Otherwise, the moderator had to rely on the video recording of what the participant was seeing to understand what happened.

However, the most challenging aspect of this evaluation was the multiple tasks that the moderator had to perform simultaneously. This included keeping the participants relaxed, preparing the equipment, keeping notes of what was happening during the testing session, prompting participants, paying attention to time and looking at the shared screen for additional information regarding pain points. This required a lot of concentration throughout the testing session. A few times, to respect the availabilities of participants, the moderator had to have consecutive back-to-back testing sessions. As such, it would have been beneficial to have another moderator to assist with the testing process of the usability evaluation in this study.

## Questionnaires

### Comfort questionnaire

The Comfort questionnaire reported how participants felt before, and after their interaction with the headset. For some items of the questionnaire it was possible to use data from the testing session and the interviews to interpret the reasons behind the changes in condition of participants. For example, there was an increase of eye strain reported by participants after using the headset. Participants during the test explained using the think-aloud method that certain menus and logos appeared too close to their eyes in some of the videos which made them feel physically uncomfortable. Similar comments were made about the menus and logos during the interview by other participants.

Moreover, it was additionally reported that five participants overall had "Difficulty in concentrating" before the experiment. One participant told the moderator that he had a migraine before the testing session while another explained that he felt stressed on that day. In the post-comfort questionnaire, the total number of participants experiencing the condition was reduced down to three participants. Later in the interviews, the participant who had the migraine before the experiment reported that the migraine had disappeared, and the other participant reported that his “stress level was really reduced”.

Overall, the comfort questionnaire was able to describe changes that participants felt before and after their experience with the headset. Combined with the data from the interviews it was possible to get some additional insight about the reasons behind the changes.

**System Usability Scale questionnaire (SUS)**

The SUS questionnaire was used to evaluate the aspect of overall usability and satisfaction of the headset experience. A product needs to score above 68 in the SUS scoring system to be considered above average. Participants in this study graded the headset with a total score of 73. Additional insight regarding usability problems that can lower the SUS score were discovered in the interview (see APPENDIX E).

**Presence questionnaire**

The questionnaire was used to examine the degree that participants experienced presence while inside the virtual environment. The overall presence score that was given was 73.

It was interesting to observe that the category “Realism” received the highest points out of the five categories despite half of the participants reporting the graphics not being of high quality. Some of the participants in the interview mentioned that the experience did not feel real due to the “resolution” of the experience, and the “quality of the image”. Similarly, some participants observed this in the video “Phillipines underwater” which was described as having “poorer quality” and “shaky lines” compared to other videos of the experience. These observations are in agreement with previous research that describe how visual quality influences immersion in Virtual environments [Bowman and McMahan 2007; Cummings and Bailenson 2016; Slater and Wilbur 1997]. On another note, the remaining half of the participants reported that the experience felt realistic which helped them get immersed. 360 degrees videos were considered to be more engaging than their 2D counterparts.

The questionnaire was able to provide an overall score for presence and indicated that realism was considered to be the most influential aspect out of all presence categories. Comments from participants described the quality of the video and video format as factors that influenced how realism was experienced.

## Summary

I consider that the methods used in this study were able to produce satisfactory data measuring the experience with this headset. However, across the experiment and analysis of this data there were certain difficulties.

Through the organization and analysis of the collected data, I was able to identify and understand the pain points of participants' experiences and the root causes of certain issues they encountered. However, such a process was lengthy and time consuming for a single person. My recommendation would be to do this sort of organization and analysis with the help of another researcher and include fewer participants depending on which aspect of the experience is being evaluated. For example, as this study was trying to expose pain points of users, six or seven participants could have been used. This would make the overall testing process more manageable by one researcher.

Usability testing has proven to be a great fit for exposing issues in Virtual Reality evaluations. Through usability testing, the moderator was able to observe the participants interacting with the system, note their thoughts and see the results of their actions. However, while this method worked well with the given experiment, certain issues were encountered such as issues with screen-sharing the user view to another computer, devices not operating as expected, shared video from the headset appearing too small on another computer, the headset covering the face of the participant and multitasking.

One of the methods in this study was questionnaires. The questionnaires were able to provide data which were quite useful for the overall analysis. However, there were unable to explain the root causes of participants' satisfactions and dissatisfactions. To get additional insight the researcher had to utilize the interviews to interpret some of the results. Table 6 summarizes the aspects of Virtual Reality User-experience described in this study and the methods that were used to evaluate them.

| Virtual Reality User-Experience | |
|---|---|
| Aspects of Virtual Reality experience | Methods used to evaluate aspects |
| Ergonomics | Interviews |
| | Questionnaires (Comfort questionnaire in this study) |
| | Think-aloud |
| Interaction with content | Direct observation |
| | Think-aloud |
| | Interviews |
| | Questionnaires (Presence and SUS questionnaire in this study) |
| Technical qualities | Think-aloud |
| | Direct observation |
| | Interviews |
| | Questionnaires (SUS and Presence questionnaire in this study) |
| Presence | Think-aloud |
| | Interviews |
| | Questionnaires (Presence questionnaire in this study) |

*Table 6. Aspects of Virtual Reality and methods that were used to evaluate them.*

## Contributions

The contributions of this work are:

- Suggested the idea that User-Experience in Virtual Reality both includes the aspects of the experience and is influenced by them simultaneously (see figure 2, *Chapter 2: User-Experience and usability in Virtual Reality*).
- Surveyed the existing literature to compile a list of available usability evaluation methods in Virtual Reality (see table 1, *Chapter 2: measuring Usability in VR*).
- Identified aspects that influence User-Experience in Virtual Reality in this study and the methods that can be used to evaluate them.
- Presented how well the methods were able to provide relevant data based on observations and impressions from this experiment.

## Limitations

Using think aloud during the testing session is important for gathering qualitative data, however it can be harmful in the case of virtual reality (where presence and immersion is a chief objective) as it can remind the participant that they are being tested. This issue has two aspects: 1) participants have to multitask while they are thinking aloud, which can make completing tasks of the evaluation harder, 2) instructing participants to fulfill specific, pre-orchestrated tasks makes it infeasible to witness a truthfully unbiased interaction of the participant with the system that is tested.

To some extent, comfort was negatively affected by the actions of the moderator. The headset offered the possibility to adjust the tightness of the straps of the headset into two different levels that allow the headset to sit comfortably on the head of the participant. The moderator did not notice that this was possible until the last testing sessions. In future evaluations, a closer examination to the characteristics of the headset will be performed.

Another limitation in this study was that it was organized and carried out by a single researcher. As described in this thesis, the researcher had to be in charge of all the aspects of this evaluation, including the equipment, keeping notes, observing the interaction of participants with the system, troubleshooting issues and staying focused during the testing sessions. This multitasking had an unfavorable effect on the collection of data. Twice the moderator forgot to activate the time recording of certain tasks, resulting in loss of data. Having at least one more researcher to provide help in the experiment would be ideal.

This study has used multiple questionnaires to investigate comfort, satisfaction and the feeling of Presence in Virtual Reality. However, there are two more questionnaires that are actively used within the field of Human Computer Interaction. These are the Simulator Sickness Questionnaire and the NASA-TLX questionnaire. The Simulator Sickness questionnaire is used to observe whether the users of a Virtual Reality system undergo cybersickness symptoms [Bruck and Watters 2009]. The NASA-TLX questionnaire is used to assess the personal workload of users operating with an interacting system [Noyes and Bruneau 2007]. The Comfort questionnaire used within this study was provided by the company and is essentially the same with the Simulator Sickness questionnaire. The only notable difference between the two questionnaires is that certain parameters are presented in a different order. The NASA-TLX questionnaire was not used in this study as there was no need to evaluate the subjective workload of users as they interacted with the virtual reality headset.

## 6) CONCLUSION

The goal of this study was to investigate how we can perform a holistic user-experience evaluation on a Virtual Reality headset by identifying the aspects of Virtual Reality that are influential to the experience and observe how successfully various methods were able to evaluate them. As every application and experience is unique, this study had to identify the aspects that were important to this particular experience and the methods that could be used to evaluate them. The aspects of this experience were summarized as interaction with content, technical qualities of the system, presence and ergonomics.

Usability testing was successfully able to expose pain points the participants experienced related to interaction with the content and the technical qualities of the system. However, it proved to be challenging to manage for only one researcher due to the amount of multitasking that was required and unexpected technical issues. The combination of interviews, direct observation, think-aloud protocol and questionnaires collected a large amount of data regarding presence, ergonomics and pain points of this experience, but organizing and analyzing the collected information proved to be very extensive and time consuming. Questionnaires were able to provide data about the impressions of users but could not describe the reasons that created these impressions without the aid of interviews. However, despite the described challenges of each method, it was their combined usage that was effective in collecting and evaluating the data of this study.

Moreover, as the characteristics that can potentially influence the experience of a user in Virtual Reality can be numerous and overwhelming, User-Experience specialists, developers and researchers working with Virtual Reality need to identify, understand and be aware of the aspects of the technology that can potentially influence the experience of the users and the methods that can evaluate them. This is crucial as the aspects that might be considered important in one application can be irrelevant in another one. As the area of Virtual Reality technology continues to evolve rapidly, efficient yet holistic user-experience evaluations of Virtual Reality can ensure that the solutions provided meet the needs of the users and ensure a full-filling, enjoyable and pain-free experience to the users.

## REFERENCES

Abulrub, A. H. G., Attridge, A. N., & Williams, M. A. (2011). Virtual reality in engineering education: The future of creative learning. In 2011 IEEE global engineering education conference (EDUCON) (pp. 751-757). IEEE.

Aitamurto, T., Zhou, S., Sakshuwong, S., Saldivar, J., Sadeghi, Y., & Tran, A. (2018). Sense of presence, attitude change, perspective-taking and usability in first-person split-sphere 360 video. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (p. 545). ACM.

Alben, L. (1996). Quality of experience: defining the criteria for effective interaction design. *interactions*, *3*(3), 11-15.

Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., & MacIntyre, B. (2001). Recent advances in augmented reality. *IEEE computer graphics and applications*, *21*(6), 34-47.

Bell, J. T., & Fogler, H. S. (1997). Ten steps to developing virtual reality applications for engineering education. In *Proceedings of the American Society for Engineering Education Annual Conference*.

Bevan, N., Carter, J., & Harker, S. (2015). ISO 9241-11 revised: What have we learnt about usability since 1998?. In *International Conference on Human-Computer Interaction* (pp. 143-151). Springer, Cham.

Blattgerste, J., Renner, P., & Pfeiffer, T. (2018). Advantages of eye-gaze over head-gaze-based selection in virtual and augmented reality under varying field of views. In *Proceedings of the Workshop on Communication by Gaze Interaction*(p. 1). ACM.

Bowman, D. A., & Hodges, L. F. (1997). An Evaluation of Techniques for Grabbing and Manipulating Remote Objects in Immersive Virtual Environments. *SI3D*, *97*, 35-38.

Bowman, A. D. (1999). Interaction techniques for common tasks in immersive virtual environments: design, evaluation, and application [Ph. D. thesis]. *Atlanta, Ga, USA*.

Bowman, D., Kruijff, E., La Viola, J., & Poupyrev, I. (2000). I., The Art and Science of 3D Interaction, tutorial notes. In *Proceedings of IEEE International Virtual Reality 2000 conference. New Brunswick, NJ*.

Bowman, D. A., Johnson, D. B., & Hodges, L. F. (2001). Testbed evaluation of virtual environment interaction techniques. Presence: Teleoperators & Virtual Environments, 10(1), 75-95.

Bowman, D. A., Gabbard, J. L., & Hix, D. (2002). A survey of usability evaluation in virtual environments: classification and comparison of methods. *Presence: Teleoperators & Virtual Environments*, *11*(4), 404-424.

Bowman, D. A., & McMahan, R. P. (2007). Virtual reality: how much immersion is enough?. *Computer*, *40*(7), 36-43.

Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry*, *189*(194), 4-7.

Brooke, J. (2013). SUS: a retrospective. *Journal of usability studies*, *8*(2), 29-40.

Bruck, S., & Watters, P. A. (2009). Estimating cybersickness of simulated motion using the simulator sickness questionnaire (SSQ): A controlled study. In *2009 sixth international conference on computer graphics, imaging and visualization* (pp. 486-488). IEEE.

Burdea, G. C., & Coiffet, P. (2003). Virtual reality technology. John Wiley & Sons.

DIS, I. (2010). 9241-210: 2010. Ergonomics of human system interaction-Part 210: Human-centred design for interactive systems (formerly known as 13407). *International Standardization Organization (ISO). Switzerland*.

Cao, S. (2016). Virtual reality applications in rehabilitation. In *International Conference on Human-Computer Interaction* (pp. 3-10). Springer, Cham.

Card, S. K., Mackinlay, J. D., & Robertson, G. G. (1991). A morphological analysis of the design space of input devices. *ACM Transactions on Information Systems (TOIS)*, *9*(2), 99-122.

Cobb, S. V., Nichols, S., Ramsey, A., & Wilson, J. R. (1999). Virtual reality-induced symptoms and effects (VRISE). *Presence: Teleoperators & Virtual Environments*, *8*(2), 169-186.

Chen, J. L., & Stanney, K. M. (1999). A theoretical model of wayfinding in virtual environments: Proposed strategies for navigational aiding. *Presence*, *8*(6), 671-685.

Chen, W., Plancoulaine, A., Férey, N., Touraine, D., Nelson, J., & Bourdot, P. (2013). 6DoF navigation in virtual worlds: comparison of joystick-based and head-controlled paradigms. In *Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology* (pp. 111-114). ACM.

Cherniack, E. P. (2011). Not just fun and games: applications of virtual reality in the identification and rehabilitation of cognitive disorders of the elderly. *Disability and rehabilitation: Assistive technology*, *6*(4), 283-289.

Chien, Y. T., & Jenkins, J. (1994). Virtual reality assessment. *A report of the Task Group on Virtual Reality to the High Performance Computing and Communications and Information Technology (HPCCIT) Subcommittee of the Information and Communications Research and Development Committee of the National Science and Technology Council (NSTC)*.

Craig, A. B., Sherman, W. R., & Will, J. D. (2009). *Developing virtual reality applications: Foundations of effective design*. Morgan Kaufmann.

Cummings, J. J., & Bailenson, J. N. (2016). How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology*, *19*(2), 272-309.

D'Cruz, M. (1999). *Structured evaluation of training in virtual environments* (Doctoral dissertation, University of Nottingham).

Darken, R. P., & Sibert, J. L. (1993). A toolset for navigation in virtual environments. In *ACM symposium on User interface software and technology* (pp. 157-165).

Darken, R. P., & Sibert, J. L. (1996). Wayfinding strategies and behaviors in large virtual worlds. In *CHI* (Vol. 96, pp. 142-149).

Deol, K. K., Steed, A., Hand, C., Istance, H., & Tromp, J. (2000). Usability evaluation for virtual environments: methods, results and future directions (part 2). *Interfaces*, *44*, 4-7.

DiZio, P., & Lackner, J. R. (1992). Spatial orientation, adaptation, and motion sickness in real and virtual environments. *Presence: Teleoperators & Virtual Environments*, *1*(3), 319-328.

Durlach, N., Mavor, A. S., & Newby, G. B. (1996). Virtual Reality: Scientific and Technological Challenges. *Library and Information Science Research*, *18*(3), 278-280.

Ellis, S. R., & Meyer-Arendt, J. R. (1992). Pictorial communication in virtual and real environments. *Applied Optics*, *31*, 6348.

ESA, E. (2014). Essential facts about the computer and video game industry.

Feiner, S. K. (2002). Augmented reality: A new way of seeing. Scientific American, 286(4), 48-55.

Fenton, N. E., & Pfleeger, S. L. (1997). Software Metrics: A Rigorous & Practical Approach. Boston: PWS Publishing Company

Foley, J. D., Wallace, V. L., & Chan, P. (1984). The human factors of computer graphics interaction techniques. *IEEE computer Graphics and Applications*, *4*(11), 13-48.

Fontaine, G. (1992). The experience of a sense of presence in intercultural and international encounters. *Presence: Teleoperators & Virtual Environments*, *1*(4), 482-490.

Furnas, G. W. (1997). Effective view navigation. In *CHI* (Vol. 97, pp. 367-374).

Gabbard, J. L. (1997). *A taxonomy of usability characteristics in virtual environments* (Doctoral dissertation, Virginia Tech).

Gabbard, J. L., Hix, D., & Swan, J. E. (1999). User-centered design and evaluation of virtual environments. *IEEE computer Graphics and Applications*, *19*(6), 51-59.

Garcia-Betances, R. I., Jiménez-Mixco, V., Arredondo, M. T., & Cabrera-Umpiérrez, M. F. (2015). Using virtual reality for cognitive training of the elderly. *American Journal of Alzheimer's Disease & Other Dementias®*, *30*(1), 49-54.

Geszten, D., Hámornik, B. P., Komlodi, A., Hercegfi, K., Szabó, B., & Young, A. (2015). Qualitative analysis of user experience in a 3D virtual environment. In *Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community* (p. 124). American Society for Information Science.

Guttentag, D. A. (2010). Virtual reality: Applications and implications for tourism. *Tourism Management*, *31*(5), 637-651.

Gutierrez, M., Vexo, F., & Thalmann, D. (2008). *Stepping into virtual reality*. Springer Science & Business Media.

Haase, T., & Termath, W. (2015). A virtual interactive training application for supporting service technicians in the field of high voltage equipment. *Procedia Computer Science*, *77*, 207-214.

Hannaford, B., & Venema, S. (1995). Kinesthetic Displays for Remote and Virtual. *Virtual environments and advanced interface design*, 415.

Hansen, J. P., Rajanna, V., MacKenzie, I. S., & Bækgaard, P. (2018). A Fitts' law study of click and dwell interaction by gaze, head and mouse with a head-mounted display. In *Proceedings of the Workshop on Communication by Gaze Interaction* (p. 7). ACM.

Hamari, J., & Keronen, L. (2017). Why do people play games? A meta-analysis. *International Journal of Information Management*, *37*(3), 125-141.

Hale, K. S., & Stanney, K. M. (2014). *Handbook of virtual environments: Design, implementation, and applications*. CRC Press.

Hassenzahl, M. (2003). The thing and I: understanding the relationship between user and product. In Funology (pp. 31-42). Springer, Dordrecht.

Hassenzahl, M., & Tractinsky, N. (2006). User experience-a research agenda. *Behaviour & information technology*, *25*(2), 91-97.

Hix, D., & Hartson, H. R. (1993). Developing user interfaces: Ensuring usability through product & process. New York: John Wiley and Sons.

Hix, D., Swan, J. E., Gabbard, J. L., McGee, M., Durbin, J., & King, T. (1999). User-centered design and evaluation of a real-time battlefield visualization virtual environment. In *Proceedings IEEE Virtual Reality (Cat. No. 99CB36316)* (pp. 96-103). IEEE.

Holzinger, A. (2013). Human-Computer Interaction and Knowledge Discovery (HCI-KDD): What is the benefit of bringing those two fields to work together?. In *International Conference on Availability, Reliability, and Security*(pp. 319-328). Springer, Berlin, Heidelberg.

Hua, H. (2017). Enabling focus cues in head-mounted displays. *Proceedings of the IEEE*, *105*(5), 805-824.

Jalaliniya, S., Mardanbegi, D., & Pederson, T. (2015). MAGIC pointing for eyewear computers. In *Proceedings of the 2015 ACM International Symposium on Wearable Computers*(pp. 155-158). ACM.

Johnson, C. (1999). Evaluating the Contribution of DesktopVR for Safety—Critical Applications. In *International Conference on Computer Safety, Reliability, and Security* (pp. 67-78). Springer, Berlin, Heidelberg.

Jordan, P. W. (2002). *Designing pleasurable products: An introduction to the new human factors*. CRC press.

Johnson, J. (2013). *Designing with the mind in mind: simple guide to understanding user interface design guidelines*. Elsevier.

Kaasinen, E., Roto, V., Hakulinen, J., Heimonen, T., Jokinen, J. P., Karvonen, H., ... & Tokkonen, H. (2015). Defining user experience goals to guide the design of industrial systems. *Behaviour & Information Technology*, *34*(10), 976-991.

Khademi, M., Mousavi Hondori, H., McKenzie, A., Dodakian, L., Lopes, C. V., & Cramer, S. C. (2014). Free-hand interaction with leap motion controller for stroke rehabilitation. In *Proceedings of the extended abstracts of the 32nd annual ACM conference on Human factors in computing systems* (pp. 1663-1668). ACM.

Kan, H. Y., Duffy, V. G., & Su, C. J. (2001). An Internet virtual reality collaborative environment for effective product design. *Computers in Industry*, *45*(2), 197-213.

Kaur, K., Maiden, N., & Sutcliffe, A. (1996). Design practice and usability problems with virtual environments. *Proceedings of Virtual Reality World*, *96*.

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, *3*(3), 203-220.

Kalawsky, R. S. (1993). The science of Virtual reality and Virtual Environments, A technical. *Science and Engineering Reference on Virtual Environments*.

Kalawsky, R. S. (1999). VRUSE—a computerised diagnostic tool: for usability evaluation of virtual/synthetic environment systems. *Applied ergonomics*, *30*(1), 11-25.

Kolasinski, E. M. (1997). Prediction of simulator sickness in a virtual environment.

Kim, B. (2016). Virtual reality as an artistic medium: A study on creative projects using contemporary head-mounted displays.

Kuutti, K. (1996). Activity theory as a potential framework for human-computer interaction research. *Context and consciousness: Activity theory and human-computer interaction*, *1744*.

Law, E., Scapin, D., Cockton, G., Stary, M., & Winckler, M. (2009). Maturation of usability evaluation methods: Retrospect and prospect. In *COST294-MAUSE Closing Conference Proceedings(2009)*.

Lallemand, C., Gronier, G., & Koenig, V. (2015). User experience: A concept without consensus? Exploring practitioners' perspectives through an international survey. *Computers in Human Behavior*, *43*, 35-48.

Li, L., Yu, F., Shi, D., Shi, J., Tian, Z., Yang, J., ... & Jiang, Q. (2017). Application of virtual reality technology in clinical medicine. *American journal of translational research*, *9*(9), 3867.

Maguire, M. (2013). Using human factors standards to support user experience and agile design. In *International Conference on Universal Access in Human-Computer Interaction* (pp. 185-194). Springer, Berlin, Heidelberg.

Mäkelä, A., & Fulton Suri, J. (2001). Supporting users' creativity: Design to induce pleasurable experiences. In *Proceedings of the International Conference on Affective Human Factors Design* (pp. 387-394).

Minakata, K., Hansen, J. P., Bækgaard, P., MacKenzie, I. S., & Rajanna, V. (2019). Pointing by gaze, head, and foot in a head-mounted display. In *ETRA´ 19*.

Mine, M. R. (1995). Virtual environment interaction techniques. *UNC Chapel Hill CS Dept*.

Mitsuhara, H., Iguchi, K., & Shishibori, M. (2017). Using Digital Game, Augmented Reality, and Head Mounted Displays for Immediate-Action Commander Training. *International Journal of Emerging Technologies in Learning*, *12*(2).

Montero, F., Lozano, M. D., & González, P. (2008). Usability-oriented quality model based on ergonomic criteria. In *Handbook of research on web information systems quality* (pp. 220-233). IGI Global.

Muhanna, M. A. (2015). Virtual reality and the CAVE: Taxonomy, interaction challenges and research directions. *Journal of King Saud University-Computer and Information Sciences*, *27*(3), 344-361.

Murphy, H., & Duchowski, A. T. (2001). Gaze-contingent level of detail rendering. *EuroGraphics 2001*.

Neale, H., & Nichols, S. (2001). Theme-based content analysis: a flexible method for virtual environment evaluation. *International journal of human-computer studies*, *55*(2), 167-189.

Nielsen, J., & Landauer, T. K. (1993). A mathematical model of the finding of usability problems. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems* (pp. 206-213). ACM.

Nielsen, J. (1994). *Usability engineering*. Elsevier.

Nielsen, J., & Mack, R. L. (1994b). Executive summary. In J. Nielsen & R. L. Mack (Eds.), Usability Inspection Methods. (pp. 1–23). New York: John Wiley & Sons

Nichols, S., Cobb, S., & Wilson, J. R. (1997). Health and safety implications of virtual environments: Measurement issues. *Presence: Teleoperators & Virtual Environments*, *6*(6), 667-675.

Nichols, S. (1998). Development of a questionnaire to measure attitudes towards virtual reality. *Contemporary Ergonomics*, 146-150.

Nichols, S. (1999). *Virtual Reality Induced Symptoms and Effects (VRISE): methodological and theoretical issues* (Doctoral dissertation, University of Nottingham).

Nichols, S., Haldane, C., & Wilson, J. R. (2000). Measurement of presence and its consequences in virtual environments. *International Journal of Human-Computer Studies*, *52*(3), 471-491.

Norman, D. A., & Draper, S. W. (1986). *User centered system design: New perspectives on human-computer interaction*. CRC Press.

Norman, D. A. (2004). *Emotional design: Why we love (or hate) everyday things*. Basic Civitas Books.

Norman, D. A. (2005). Human-centered design considered harmful. *interactions*, *12*(4), 14-19.

Noyes, J. M., & Bruneau, D. P. (2007). A self-analysis of the NASA-TLX workload measure. *Ergonomics, 50*(4), 514-519.

Obrist, M., Roto, V., & Väänänen-Vainio-Mattila, K. (2009). User experience evaluation: do you know which method to use?. In *CHI'09 Extended Abstracts on Human Factors in Computing Systems* (pp. 2763-2766). ACM.

Paul, R. P. (1981). *Robot manipulators: mathematics, programming, and control: the computer control of robot manipulators*. Richard Paul.

Pausch, R., Shackelford, M.A., Proffitt, D., 1993. A user study comparing head-mounted and stationary displays. In: 1993 IEEE Symposium on Research Frontiers in Virtual Reality, pp. 41–45.

Pierce, J. S., Forsberg, A. S., Conway, M. J., Hong, S., Zeleznik, R. C., & Mine, M. R. (1997). Image plane interaction techniques in 3D immersive environments. In *Proceedings of the 1997 symposium on Interactive 3D graphics* (pp. 39-ff). ACM.

Piumsomboon, T., Lee, G., Lindeman, R. W., & Billinghurst, M. (2017). Exploring natural eye-gaze-based interaction for immersive virtual reality. In *2017 IEEE Symposium on 3D User Interfaces (3DUI)* (pp. 36-39). IEEE.

Plunkett, L. (2014). Facebook buys oculus rift for $2 billion. *Kotaku. com*.

Polson, P. G., Lewis, C., Rieman, J., & Wharton, C. (1992). Cognitive walkthroughs: a method for theory-based evaluation of user interfaces. *International Journal of man-machine studies*, *36*(5), 741-773.

Poupyrev, I., Weghorst, S., Billinghurst, M., & Ichikawa, T. (1997). A framework and testbed for studying manipulation techniques for immersive VR. In *VRST* (Vol. 97, pp. 21-28).

Poupyrev, I., Billinghurst, M., Weghorst, S., & Ichikawa, T. (1996). The go-go interaction technique: non-linear mapping for direct manipulation in VR. In Proceedings of the ACM Symposium on User Interface Software and Technology – UIST '96, (pp. 79-80). ACM.

Poupyrev, I. (2000). Immersive techniques for virtual manipulation. In *The Art and Science of 3D Interaction, tutorial notes from the IEEE International Virtual Reality 2000 Conference, New Brunswick, NJ March* (pp. 18-22).

Psotka, J., & Davison, S. (1993). Cognitive factors associated with immersion in virtual environments.

Qian, Y. Y., & Teather, R. J. (2017). The eyes don't have it: an empirical comparison of head-based and eye-based selection in virtual reality. In *Proceedings of the 5th Symposium on Spatial User Interaction* (pp. 91-98). ACM.

Qian, L., Deguet, A., & Kazanzides, P. (2018). ARssist: augmented reality on a head-mounted display for the first assistant in robotic surgery. *Healthcare technology letters*, *5*(5), 194-200.

Regan, E. C. (1995). Some human factors issues in immersive virtual reality: Fact and speculation. In *Virtual reality applications* (pp. 163-182). Academic Press Ltd.

Reingold, E. M., Loschky, L. C., McConkie, G. W., & Stampe, D. M. (2003). Gaze-contingent multiresolutional displays: An integrative review. *Human factors*, *45*(2), 307-328.

Reason, J. T., & Brand, J. J. (1975). *Motion sickness*. Academic press.

Richard, P., Birebent, G., Coiffet, P., Burdea, G., Gomez, D., & Langrana, N. (1996). Effect of frame rate and force feedback on virtual object manipulation. *Presence: Teleoperators & Virtual Environments*, *5*(1), 95-108.

Riva, G., Waterworth, J., & Murray, D. (Eds.). (2014). *Interacting with Presence: HCI and the Sense of Presence in Computer-mediated Environments*. Walter de Gruyter GmbH & Co KG.

Rogers, Y. (2012). HCI theory: classical, modern, and contemporary. *Synthesis lectures on human-centered informatics*, *5*(2), 1-129.

Roto, V., Rantavuo, H., & Väänänen-Vainio-Mattila, K. (2009). Evaluating user experience of early product concepts. In *Proc. DPPI* (Vol. 9, pp. 199-208).

Rushton, S., Mon-Williams, M., Wann, J.P., 1994. Binocular vision in a bi-ocular world: new generation head-mounted displays avoid causing visual deficit. Displays 15 (4), 255–260.

Sanders, E. (1999). Design for experiencing: New tools. *Proc. Design and Emotion, 1999*.

Salvendy, G. (Ed.). (2012). *Handbook of human factors and ergonomics*. John Wiley & Sons.

Scriven, M. (1967). The Methodology of Evaluation. Perspectives of Curriculum Evaluation. AERA. Monograph 1.

Schofield, D. (2012). Mass effect: A chemical engineering education application of virtual reality simulator technology. *Journal of Online Learning and Teaching*, *8*(1), 63.

Seffah, A., & Taleb, M. (2012). Tracing the evolution of HCI patterns as an interaction design tool. *Innovations in Systems and Software Engineering*, *8*(2), 93-109.

Sharp, H., Rogers, Y., & Preece, J. (2007). Interaction Design: Beyond Human-Computer Interaction (pp. 1-551). Chichester, England: John Wiley & Sons.

Sharples, S., Stedmon, A. W., D'Cruz, M., Patel, H., Cobb, S., Yates, T., ... & Wilson, J. R. (2007). Human Factors of Virtual Reality–Where are We Now?. In *Meeting diversity in ergonomics*(pp. 173-186). Elsevier Science Amsterdam.

Slater, M., Usoh, M., & Steed, A. (1994). Depth of presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, *3*(2), 130-144.

Slater, M., & Usoh, M. (1994). Body centred interaction in immersive virtual environments. *Artificial life and virtual reality*, *1*(1994), 125-148.

Slater, M., Usoh, M., & Steed, A. (1995). Taking steps: the influence of a walking technique on presence in virtual reality. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *2*(3), 201-219.

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, *6*(6), 603-616.

Siu, K. C., Best, B. J., Kim, J. W., Oleynikov, D., & Ritter, F. E. (2016). Adaptive virtual reality training to optimize military medical skills acquisition and retention. *Military medicine*, *181*(suppl_5), 214-220.

Smith, M. J., Ginger, E. J., Wright, K., Wright, M. A., Taylor, J. L., Humm, L. B., ... & Fleming, M. F. (2014). Virtual reality job interview training in adults with autism spectrum disorder. *Journal of autism and developmental disorders*, *44*(10), 2450-2463.

Soukoreff, R. W., & MacKenzie, I. S. (2004). Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI. *International journal of human-computer studies*, *61*(6), 751-789.

Stanney, K., & Salvendy, G. (1998). Aftereffects and sense of presence in virtual environments: Formulation of a research and development agenda. *International Journal of Human-Computer Interaction*, *10*(2), 135-187.

Stanney, K. M., & Reeves, L. (2000). COVE evaluation report. *Final Report, Contract*, (61339).

Stanney, K. M., Mollaghasemi, M., Reeves, L., Breaux, R., & Graeber, D. A. (2003). Usability engineering of virtual environments (VEs): identifying multiple criteria that drive effective VE system design. *International Journal of Human-Computer Studies*, *58*(4), 447-481.

Stanica, I. C., Dascalu, M. I., Moldoveanu, A., Bodea, C. N., & Hostiuc, S. (2016). A survey of virtual reality applications as psychotherapeutic tools to treat phobias. In *The International*

*Scientific Conference eLearning and Software for Education* (Vol. 1, p. 392). " Carol I" National Defence University.

Steed, A., & Tromp, J. (1998, June). Experiences with the Evaluation of CVE Applications. In *Proceedings of collaborative virtual environments* (pp. 123-127).

Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of communication*, *42*(4), 73-93.

Steven, M. L. (2016). Virtual Reality.

Stone, R. J. (2011). The (human) science of medical virtual learning environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1562), 276-285.

Teather, R. J., & Stuerzlinger, W. (2013, April). Pointing at 3d target projections with one-eyed and stereo cursors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 159-168). ACM.

Uvais, N. (2015). Virtual reality applications in Schizophrenia. *Industrial psychiatry journal*, *24*(2).

Väänänen-Vainio-Mattila, K., Roto, V., & Hassenzahl, M. (2008). Towards practical user experience evaluation methods. *Meaningful measures: Valid useful user experience measurement (VUUM)*, 19-22.

Van Kerrebroeck, H., Brengman, M., & Willems, K. (2017). When brands come to life: experimental research on the vividness effect of Virtual Reality in transformational marketing communications. *Virtual Reality*, *21*(4), 177-191.

Vermeeren, A. P., Law, E. L. C., Roto, V., Obrist, M., Hoonhout, J., & Väänänen-Vainio-Mattila, K. (2010). User experience evaluation methods: current state and development needs. In *Proceedings of the 6th Nordic conference on human-computer interaction: Extending boundaries* (pp. 521-530). ACM.

Wann, J. P., & Mon-Williams, M. (2002). Measurement of visual aftereffects following virtual environment exposure. In *Handbook of virtual environments* (pp. 771-790). CRC Press.

Ware, C., & Balakrishnan, R. (1994). Reaching for objects in VR displays: lag and frame rate. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *1*(4), 331-356.

Wickens, C. D., & Baker, P. (1995). Cognitive issues in virtual reality.

Wilson, M. S., Middlebrook, A., Sutton, C., Stone, R., & McCloy, R. F. (1997). MIST VR: a virtual reality trainer for laparoscopic surgery assesses performance. *Annals of the Royal College of Surgeons of England*, *79*(6), 403.

Witmer, B. G., & Singer, M. J. (1994). *Measuring immersion in virtual environments*. Tech. Rep. 1014, US Army Research Institute for the Behavioral and Social Sciences.

Winn, W. (2002). Research into practice: Current trends in educational technology research: The study of learning environments. *Educational psychology review*, *14*(3), 331-351.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence*, *7*(3), 225-240.

Wright, P., McCarthy, J., & Meekison, L. (2005). Making sense of experience. Funology.

Zeltzer, D. (1992). Autonomy, interaction, and presence. *Presence: Teleoperators & Virtual Environments*, *1*(1), 127-132.

Zhai, S., Milgram, P., & Buxton, W. (1996, April). The influence of muscle groups on performance of multiple degree-of-freedom input. In *CHI* (Vol. 96, pp. 308-315).

Zimmermann, Philippe Georges. *Beyond usability: measuring aspects of user experience*. Diss. ETH Zurich, 2008.

# APPENDICES

## APPENDIX A: PRE-EXPERIMENT QUESTIONNAIRE

Participant ID: *

Your answer

Age: *

Your answer

Gender *

- ○ Male
- ○ Female
- ○ Other

Please fill your level of interest in Virtual Reality: *

- ○ No interest at all
- ○ Little interest
- ○ Neutral interest
- ○ Quite interested
- ○ Really interested

Please explain the answer to the previous question: *

Your answer

Which aspects of Virtual Reality are you most interested in?

Your answer

Have you had any previous experience with Virtual Reality headsets? *

- ◯ Yes
- ◯ No

Are you wearing glasses today? *

- ◯ Yes
- ◯ No

How many hours do you use computers, smartphones, tablets during the day? *

- ◯ 0 - 2 hours
- ◯ 3 - 4 hours
- ◯ 4 - 8 hours
- ◯ More than 8 hours

Do you have any physical or cognitive impairments that may possibly stop you from participating in this study? *

- ◯ Yes
- ◯ No

If you answered yes, please let me know of physical or cognitive impairments that might prevent you from completing this study:

Your answer

NEXT

**APPENDIX B: COMFORT QUESTIONNAIRE**

Select how much each symptom below is affecting you right now:

| | None | Slight | Moderate | Severe |
|---|---|---|---|---|
| 1. General Discomfort | ○ | ○ | ○ | ○ |
| 2. Fatigue | ○ | ○ | ○ | ○ |
| 3. Headache | ○ | ○ | ○ | ○ |
| 4. Eye strain | ○ | ○ | ○ | ○ |
| 5. Difficulty focusing | ○ | ○ | ○ | ○ |
| 6. Salivation increasing | ○ | ○ | ○ | ○ |
| 7. Sweating | ○ | ○ | ○ | ○ |
| 8. Nausea | ○ | ○ | ○ | ○ |
| 9. Difficulty concentrating | ○ | ○ | ○ | ○ |
| 10. "Fullness of the head" | ○ | ○ | ○ | ○ |
| 11. Blurred Vision | ○ | ○ | ○ | ○ |
| 12. Dizziness with open eyes | ○ | ○ | ○ | ○ |
| 13. Dizziness with closed eyes | ○ | ○ | ○ | ○ |
| 14. Vertigo | ○ | ○ | ○ | ○ |
| 15. Stomach awareness | ○ | ○ | ○ | ○ |
| 16. Burping | ○ | ○ | ○ | ○ |

## APPENDIX C: SYSTEM USABILITY SCALE QUESTIONNAIRE

For each of the following statements, mark one circle that best describes your reaction to the experience today.

| | Strongly Disagree | Disagree | Neither Agree nor Disagree | Agree | Strongly Agree |
|---|---|---|---|---|---|
| I think that I would like to use this headset frequently | ○ | ○ | ○ | ○ | ○ |
| I found this headset unnecessarily complex | ○ | ○ | ○ | ○ | ○ |
| I thought this headset was easy to use | ○ | ○ | ○ | ○ | ○ |
| I think that I would need assistance to be able to use this VR experience | ○ | ○ | ○ | ○ | ○ |
| I found the various functions in this headset were well integrated | ○ | ○ | ○ | ○ | ○ |
| I thought there was too much inconsistency in the headset | ○ | ○ | ○ | ○ | ○ |
| I would imagine that most people would learn to use this headset very quickly | ○ | ○ | ○ | ○ | ○ |
| I found the headset very cumbersome/awkward to use | ○ | ○ | ○ | ○ | ○ |
| I feel confident using the headset | ○ | ○ | ○ | ○ | ○ |
| I needed to learn a lot of things before I could get going with this headset | ○ | ○ | ○ | ○ | ○ |

# APPENDIX D: PRESENCE QUESTIONNAIRE

How much were you able to control the events?

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Not at all | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Completely |

How responsive was the environment to actions that you initiated (or performed?)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Not responsive | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Completely responsive |

How natural did your interactions with the environment seem?

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Extremely artificial | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Completely natural |

How much did the visual aspects of the environment involve you?

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Not at all | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Completely |

How natural was the mechanism which controlled changing between the different environments?

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Extremely artificial | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Completely natural |

How much did your experiences in the virtual environment seem consistent with your real world experiences?

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Not consistent | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Very consistent |

Were you able to anticipate what would happen next in response to the actions that you performed?

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Not at all | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Completely |

How completely were you able to actively survey or search the environment using your vision?

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Not at all | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Completely |

How closely were you able to examine elements in the experience?

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Not at all | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Very closely |

How involved were you in the virtual environment experience?

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Not involved | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Completely engrossed |

How much delay did you experience between your actions and expected outcomes?

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Long delays | ○ | ○ | ○ | ○ | ○ | ○ | ○ | No delays |

How quickly did you adjust to the virtual environment experience?

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Not at all | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Less than one minute |

How proficient in moving and interacting with the virtual environment did you feel at the end of experience?

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Not proficient | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Very proficient |

How much did the visual display quality interfere or distract you from performing assigned tasks or required activities?

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Prevented task performance | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Not at all |

How much did the control devices interfere with the performance of assigned tasks or with other activities?

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|

How well could you concentrate on the assigned tasks or required activities rather than on the mechanisms used to perform those tasks or activities?

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Not at all | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Completely |

## APPENDIX E: FOLLOW-UP INTERVIEW QUESTIONS

1) How was the experience of using the virtual reality headset?
Would you use it again?
Could you describe how it felt wearing the headset with a few words?

2) Is there anything that caught your eye specifically during the experience?
Something which you haven't experienced before?

3) What were 3 things you liked most about the experience?
What were 3 things you did not like?

4) If you could change something in the experience, what would that be?

5) Did you find the instructions of the tutorial clear and easy to follow?
Was something confusing you?

6) Do you feel the information of the tutorial assisted you to understand how the headset works?

7) I would like you to think of the tutorial for a moment:
How did you find the visual look of the tutorial overall?
How do you like it as a first part of the experience?

8) How did following the tutorial instructions felt in the beginning of our session?

9) Would you change something from the form of the current tutorial?
(If they struggle: is there anything that you think that was missing from the tutorial?)

10) How did it feel navigating through the content of the headset?
Did you encounter any difficulties?

11) Was the cube menu easy to locate and notice?

Was the information on the quick bar menu easy to understand?
Was the cube easy to find when the video played?

12) How did it feel to select information with the headset?
Which was your preferable method of selecting and why?

13) I would like you to think about when you watched one of the videos.
Were the video controls easy to discover?
Was the information displayed on the video player in a comprehensive way?
Was the information enough?
Would you change something related to the video player?

14) Do you have any other comments that you would like to share with me?

## APPENDIX F: CONSENT FORM

CONSENT TO RECORD A USABILITY TEST

I am asking you to participate in a usability test that is part of my thesis project work. Joining this study is voluntary. By participating in this usability test you will be helping me to evaluate the usability and holistic user-experience of a VR headset. The software of this headset has been developed by a specific company whose name I am not allowed to mention here due to a confidentiality agreement. In collaboration with them, I was enlisted with the task of conducting this test session.

During the test you will be asked to perform tasks using the VR headset and to think-aloud while doing so. In addition, after completing the tasks, I will ask you to fill in some questionnaires and interview you regarding your experience with the product. The whole duration of the experiment will be around one hour maximum.

There have been cases where human interaction with virtual reality resulted in queasiness, migraine, eye tiredness and dizziness. Regardless, you can stop your participation at any moment if you experience any of the above symptoms or for any other reason by letting me know.

During the test, I will be recording what you see from the headset screen on my laptop and its events as well as audio of what is happening in the room. The materials recorded during the test will be used to evaluate the usability and user-experience of the product for my thesis. The recordings will be shared with the company for their internal use, but your identity will remain confidential.

The results of the test will be reported anonymously and no correlation between you and the data shall be revealed. In order to do that, during the whole session you will be assigned a Participant ID which you will use to fill in the questionnaires and answer to the interviews. In that way your identify will additionally remain hidden after the testing session is over.

You can stop participating in the usability test at any point.
At this point, I will be happy to answer any questions you might have.

For the participant:

I understood the aforementioned information and I accept to take part in this testing session.

Date and place: ______________________________________________

Name: ______________________________________________

## APPENDIX G: OBSERVATION CHECKLIST

| | Comments: | Major problems with task | Assistance required by moderator: | Time to complete task: | Success Rate: |
|---|---|---|---|---|---|
| Task 1: | | | | | |
| Task 2: | | | | | |
| Task 3: | | | | | |
| Task 4: | | | | | |
| Task 5: | | | | | |
| Task 6: | | | | | |
| Task 7: | | | | | |
| Task 8: | | | | | |
| Task 9: | | | | | |
| Task 10: | | | | | |
| Task 11: | | | | | |
| Task 12: | | | | | |
| Task 13: | | | | | |
| Task 14: | | | | | |
| Task 15: | | | | | |
| Task 16: | | | | | |
| Task 17: | | | | | |
| Task 18: | | | | | |
| Task 19: | | | | | |
| Task 20: | | | | | |

## APPENDIX H: MATRIX TABLE

This is an example from the matrix table used to provide guidance on where to locate information related to the breadcrumbs task for this study.

| | |
|---|---|
| **Do participants understand the breadcrumbs?**<br><br>Summary: | - Check Time expired until successful completion of task 8:<br>- Check if task 7, 8, 9 was successful:<br>- Check whether prompting or assistance was required from checklist<br>- Check major problems associated with task from checklist<br>- Observations/comments about actions of user<br>- Count error rates of participants<br>- Observe data collected from interview question 10 |

## APPENDIX I: RAW DATA

**System Usability Scale (SUS)**

The following table presents the SUS scores for all of the participants and the overall average SUS score (n=14). Additionally, the table presents the SUS score for participants the participants that had prior Virtual Reality experience (n=7), and the SUS score for participants that did not have prior Virtual Reality experience (n=7).

*Explains which participants reported having previous experience with Virtual Reality (it could potentially be an experience that was different from the headset that was being tested)

| | SUS SCORE |
|---|---|
| Pilot participant | *75 |
| Participant 1 | *70 |
| Participant 2 | *87.5 |
| Participant 3 | *62.5 |
| Participant 4 | 67.5 |
| Participant 5 | 72.5 |
| Participant 6 | 70 |
| Participant 7 | 85 |
| Participant 8 | *65 |
| Participant 9 | 67.5 |
| Participant 10 | 85 |
| Participant 11 | *85 |
| Participant 12 | 47.5 |

| | |
|---|---|
| Participant 13 | *85 |
| **OVERALL AVERAGE** SUS SCORE: | **73.21428571** |
| *Average score of participants **with prior** Virtual Reality experience: | **76** |
| Average score of participants **without prior Virtual Reality experience:** | **71** |

**Comfort questionnaire results**

The numbers in the tables below present how many participants reported feeling a specific symptom before and after the experiment took place. Responses were given on a scale from NONE to SEVERE. The two columns named "before" and "after" present the overall results before and after the experience next to each other.

**GENERAL DISCOMFORT**

| | BEFORE | AFTER |
|---|---|---|
| NONE | 12 (85%) | 7 (50%) |
| SLIGHT | 2 (15%) | 5 (35%) |
| MODERATE | 0 | 2 (15%) |
| SEVERE | 0 | 0 |

**FATIGUE**

| | BEFORE | | AFTER | |
|---|---|---|---|---|
| NONE | 11 | (80%) | 10 | (72%) |
| SLIGHT | 3 | (20%) | 4 | (28%) |
| MODERATE | 0 | | 0 | |
| SEVERE | 0 | | 0 | |

**HEADACHE**

| | BEFORE | | AFTER | |
|---|---|---|---|---|
| NONE | 13 | (93%) | 12 | (86%) |
| SLIGHT | 1 | (7%) | 2 | (14%) |
| MODERATE | 0 | | 0 | |
| SEVERE | 0 | | 0 | |

**EYE STRAIN**

| | BEFORE | | AFTER | |
|---|---|---|---|---|
| NONE | 11 | (79%) | 6 | (43%) |
| SLIGHT | 3 | (21%) | 4 | (29%) |
| MODERATE | 0 | | 3 | (21%) |

| | | |
|---|---|---|
| SEVERE | 0 | 1 (7%) |

**DIFFICULTY FOCUSING**

| | BEFORE | AFTER |
|---|---|---|
| NONE | 9 (65%) | 9 (65%) |
| SLIGHT | 4 (29%) | 3 (21%) |
| MODERATE | 1 (6%) | 2 (14%) |
| SEVERE | 0 | 0 |

**SALIVATION INCREASING**

| | BEFORE | AFTER |
|---|---|---|
| NONE | 12 (86%) | 12 (86%) |
| SLIGHT | 2 (14%) | 2 (14%) |
| MODERATE | 0 | 0 |
| SEVERE | 0 | 0 |

**SWEATING**

| | BEFORE | AFTER |
|---|---|---|
| NONE | 13 (93%) | 10 (72%) |
| SLIGHT | 1 (7%) | 3 (22%) |
| MODERATE | 0 | 1 (6%) |
| SEVERE | 0 | 0 |

**NAUSEA**

| | BEFORE | AFTER |
|---|---|---|
| NONE | 14 (100%) | 9 (65%) |
| SLIGHT | 0 | 5 (35%) |
| MODERATE | 0 | 0 |
| SEVERE | 0 | 0 |

**DIFFICULTY CONCENTRATING**

| | BEFORE | AFTER |
|---|---|---|
| NONE | 9 (65%) | 10 (73%) |
| SLIGHT | 4 (29%) | 3 (27%) |
| MODERATE | 1 (6%) | 0 |

| | | |
|---|---|---|
| SEVERE | 0 | 0 |

**'FULNESS OF THE HEAD'**

| | BEFORE | AFTER |
|---|---|---|
| NONE | 8 (57%) | 9 (65%) |
| SLIGHT | 4 (29%) | 2 (14%) |
| MODERATE | 1 (7%) | 2 (14%) |
| SEVERE | 1 (7%) | 1 (7%) |

**BLURRED VISION**

| | BEFORE | AFTER |
|---|---|---|
| NONE | 12 (86%) | 9 (65%) |
| SLIGHT | 2 (14%) | 4 (29%) |
| MODERATE | 0 | 1 (6%) |
| SEVERE | 0 | 0 |

**DIZZINESS WITH OPEN EYES**

| | BEFORE | AFTER |
|---|---|---|
| NONE | 13 (93%) | 9 (65%) |
| SLIGHT | 1 (7%) | 4 (28%) |
| MODERATE | 0 | 1 (7%) |
| SEVERE | 0 | 0 |

**DIZZINESS WITH CLOSED EYES**

| | BEFORE | AFTER |
|---|---|---|
| NONE | 14 (100%) | 10 (72%) |
| SLIGHT | 0 | 3 (21%) |
| MODERATE | 0 | 1 (7%) |
| SEVERE | 0 | 0 |

**VERTIGO**

| | BEFORE | AFTER |
|---|---|---|
| NONE | 14 (100%) | 11 (79%) |
| SLIGHT | 0 | 3 (21%) |
| MODERATE | 0 | 0 |

| | | |
|---|---|---|
| SEVERE | 0 | 0 |

**STOMACH AWARENESS**

| | BEFORE | AFTER |
|---|---|---|
| NONE | 13 (100%) | 11 (79%) |
| SLIGHT | 1 | 3 (21%) |
| MODERATE | 0 | 0 |
| SEVERE | 0 | 0 |

**BURPING**

| | BEFORE | AFTER |
|---|---|---|
| NONE | 14 (100%) | 13 (93%) |
| SLIGHT | 0 | 1 (7%) |
| MODERATE | 0 | 0 |
| SEVERE | 0 | 0 |