Marianne Hernholm

A virtual reality pose estimation exercise game for post-stroke upper-limb motor function rehabilitation

Hovedoppgave i Datateknologi
Veileder: Xiaomeng Su
Medveileder: Emanuel Alexander Lorenz
Juni 2023

Hovedoppgave

NTNU
Norges teknisk-naturvitenskapelige universitet
Fakultet for informasjonsteknologi og elektroteknikk
Institutt for datateknologi og informatikk

NTNU
Kunnskap for en bedre verden

Marianne Hernholm

A virtual reality pose estimation exercise game for post-stroke upper-limb motor function rehabilitation

Hovedoppgave i Datateknologi
Veileder: Xiaomeng Su
Medveileder: Emanuel Alexander Lorenz
Juni 2023

Norges teknisk-naturvitenskapelige universitet
Fakultet for informasjonsteknologi og elektroteknikk
Institutt for datateknologi og informatikk

NTNU
Kunnskap for en bedre verden

Preface

This project is a master thesis in M.Sc. Computer Science at the Norwegian University of Science and Technology. A preliminary literature review was carried out in the autumn semester 2022, and the thesis work was carried out in the spring semester 2023.

The work was supervised by Xiaomeng Su and co-supervised by Emanuel Alexander Lorenz. I would like to thank my supervisors for all the motivation and help I have received. Working on this project has been very interesting and challenging, and I have learned a lot. I would also like to thank my fellow computer science students for all the motivation and encouragement along the way.

Abstract

Stroke is a leading causes of acquired disability in adults worldwide. Currently, the demand for post-stroke rehabilitation outweighs the amount of stroke care providers, and this resource shortage is only predicted to increase as life expectancy globally continues to rise. Virtual reality exercise games can provide conditions to induce neuroplasticity, a key aspect of stroke rehabilitation. The use of virtual reality interventions for post-stroke rehabilitation have gained significant footing in the healthcare field. These exercise games require a method of motion capture. Often, motion capture in virtual reality systems is achieved with methods that require additional motion capture hardware like the Kinect or marker-based solutions. An alternative is 3D human pose estimation models. They offer a method of motion capture that does not require any additional motion capture hardware, but can be used with only a digital camera and a computer. In this thesis, a novel concept for a rehabilitation game, using pose estimation for motion capture, was designed and implemented. A proof-of-concept prototype, utilizing the 2D/ 3D human pose estimation model BlazePose was developed. A usability test was conducted, on 11 test participants without motor impairments, using the standardized system usability scale. The prototype was able to garner an acceptable usability score of 83,2. Test participants identified issues regarding the control of their virtual reality avatar, the origin of which should be explored further. There is much potential for pose estimation models in the context of rehabilitation, and their application in this field should be researched further.

List of Abbreviations

ADL	Activities of Daily Living
XR	Extended Reality
HMD	Head Mounted Display
HCI	Human-Computer Interaction
HPE	Human Pose Estimation
HRI	Human-Robot Interaction
ML	Machine Learning
MR	Mixed Reality
PE	Pose Estimation
RoI	Region of Interest
SUS	System Usability Scale
UL	Upper-limb
VR	Virtual Reality
WHO	World Health Organization

Table of Contents

List of Figures **vi**

1 Introduction **1**
1.1 Motivation . . . 1
1.2 Research question and goal . . . 2
1.3 Project structure . . . 3

2 Background **4**
2.1 Stroke and motor rehabilitation . . . 4
2.2 Exercise games . . . 5
2.3 Virtual reality . . . 5
2.3.1 Immersion and presence . . . 5
2.3.2 Cybersickness . . . 5
2.4 Virtual reality in rehabilitation . . . 5
2.5 Human pose estimation . . . 6
2.5.1 BlazePose . . . 6
2.6 Development tools . . . 7
2.6.1 Unity 3D . . . 7
2.6.2 HTC Vive . . . 8
2.6.3 Webcam . . . 8
2.7 Preliminary study to determine high-level requirements . . . 9

3 Related Work **10**

4 Methods **12**
4.1 Game design process . . . 12
4.2 Exercise definition . . . 13
4.3 Concept ideation . . . 14
4.4 Virtualization . . . 14
4.5 Build . . . 14
4.5.1 Gameplay . . . 15
4.5.2 Environment . . . 15
4.5.3 BlazePose . . . 16
4.5.4 Humanoid avatar . . . 17
4.5.5 Virtual reality . . . 18
4.5.6 Setup . . . 18
4.6 Validate . . . 18

5 Results **20**
5.1 Game concept . . . 20
5.2 Exergame setup . . . 20
5.3 System implementation prototype . . . 22

6 Evaluation **24**

6.1 The System Usability Scale 24
6.2 Participants 25
6.3 Testing procedure 25
6.4 Evaluation results 26
6.4.1 Quantitative results 26
6.4.2 Qualitative results 27

7 Discussions **29**
7.1 Design approach 29
7.1.1 Choice of tools 29
7.2 Human pose estimation with BlazePose 29
7.3 Limitations 30
7.4 Evaluation 30
7.5 Further work 31

8 Conclusions **32**

References **33**

A Information and consent form for test participants **37**
A.1 Information form 37
A.2 Consent form 38

B System Usability Scale **39**

C The VR Game Canvas **40**

D Generated game concepts **41**
D.1 Maestro concept 41
D.2 Volleyball concept 41

List of Figures

2.1 The 33 Pose Landmarks of BlazePose [55] 7
2.2 Overview of BlazePose's inference pipeline [41]. Frame #1 estimates the subject's RoI. The next frames subsequently utilizes the predictions of the previous frames in order to improve pose tracking [34]. 7
2.3 The HTC Vive headset [9]. 8
2.4 The Microsoft LifeCam webcam [11]. 8

4.1 Framework of Støen and Fridheim [59] for developing serious games for rehabilitation. 13
4.2 A selection of standard upper-limb exercises for motor rehabilitation of stroke survivors taken from the STREAM evaluation tool. Exercises are shown with their difficulty logit measures and the associated standard error (SE) Hsueh et al. [28]. . 14
4.3 Low-poly fruits GameObject Prefabs 15
4.4 Scene- and inspector view of the strawberry prefab 16
4.5 BlazePose running in Unity on two different input sources. 17
4.6 Scene view of the low-poly skeleton avatar in T-pose 18

5.1 Fruit catcher game concept generated through the Støen and Fridheim [59] framework. All images accessed through creative commons license. 21
5.2 Result of the virtualization analysis of fruit catcher Støen and Fridheim [59]. . . . 22
5.3 The final technical setup for the exergame. 22
5.4 Forward position of the player: Hierarchy and scene view 23
5.5 Forward position of the player: Scene view 23

6.1 Adjective rating scale for SUS scores [2] 25
6.2 Average score (original scale) for each of the 10 SUS statements with belonging standard deviation. 26
6.3 The individual SUS score for each of the 11 test participants. 27
6.4 Aspects of the system addressed by participants in the written feedback 28

C.1 The VR Game Canvas [59] 40

D.1 Maestro game concept generated through the Støen and Fridheim [59] framework. All images accessed through creative commons license. 42
D.2 Volleyball game concept generated through the Støen and Fridheim [59] framework. All images accessed through creative commons license. 43

Chapter 1

Introduction

1.1 Motivation

Today, the global demand for rehabilitation is not sufficiently met [50]. An identified reason behind the inability to match the current global demand is that there is simply a scarcity of available rehabilitation services [50]. In addition to the current demand going unmet, there is, worldwide, a predicted increase in demand for the coming years [21]. A key reason behind this increase in demand for rehabilitation is the decreased fatality rates for chronic conditions such as stroke and cancer, due to medical advancements [50, 21, 16]. A level of acquired disability is a common consequence for survivors of these conditions. Rehabilitation is a highly beneficial tool for lessening the burden of these negative, acquired symptoms [42].

Stroke is one of the leading causes of acquired disability in adults worldwide [57]. The risk of experiencing a stroke increases with age, meaning more and more people will have strokes as the life expectancy around the world continues to increase [51, 57]. From 2015 to 2050 the percentage of people over the age of 60 in the total population is expected to have doubled from 12% to 22%, landing at approximately 2.1 billion people [49, 39]. Between 2020 and 2050 the number of people aged 80 years or above is expected to triple, landing at around 426 million people [49]. Approximately 1 in 4 men and 1 in 5 women that have surpassed the age of 45 will have experienced a stroke if they live to the age of 85 [72]. Based on these predictions, one must assume the amount of people living with chronic stroke globally will only increase.

Stroke care and rehabilitation is vital for minimizing the negative effects of disability following a stroke, and improvements in these fields are essential for the continued positive advancements in the prospects of stroke victims [42]. Rehabilitation schemes that can provide increased task-specific training without increasing the workload of healthcare staff will be desired [36]. Research into the creation of systems that can counteract the shortage in providers- and resources of stroke rehabilitation [57], and contribute to meet the future demand for stroke recovery is of crucial importance in order to be able to provide future stroke victims with a dignified and decent recovery process.

Motor function rehabilitation requires the facilitation of training conditions that can induce mechanisms of neuroplasticity in the brain. Conventional therapy methods can struggle to provide optimal environments for this training [57]. Additionally, standard rehabilitation methods encounter certain limitations. Some examples are that the rehabilitation procedure is labor-and resource-intensive, geography dependent, time consuming, and that patient compliance can vary [57].

Virtual reality systems offer users in stroke recovery features that can not be replicated in conventional rehabilitation. The immersive aspect of virtual reality can in the context of rehabilitation provide av advantage as the user becomes very absorbed in the game itself. This can lead to the user focusing less on their own physical limitations, and more on the current rehabilitation scheme they are engaging with [18]. VR games have been used in many serious contexts, such as education, training, and healthcare [36]. In serious games for exercise, the users are interacting with the game using bodily movements that are captured by cameras or other devices [48]. Tracking the movements of the user so it can be utilized as input for the game requires a method of motion capture. There exist a number of technologies and methods for motion tracking that are compatible with VR interventions. Users of VR movement rehabilitation interventions for neurological conditions like stroke have reported experiencing difficulty with using hand-held motion tracking interfaces [37]. Handheld devices make an assumption regarding the functionality level of the user, namely that they do not inhabit any impairments that would make holding the device in itself challenging [23]. As hand-held interfaces can put users in motor rehabilitation with physical limitations at an initial disadvantage, their use is not applicable for this system. An alternative to hand-held interfaces are either marker-based or marker-less tracking interfaces. Marker-based motion tracking schemes are often exceedingly expensive as they require marker tools and dedicated cameras, making them lacking with regards to accessibility [58, 52]. Marker-less motion tracking interventions do not require the use of expensive, wearable sensors, and thus become the most well suited alternative for the motion tracking aspect of this system. There are a variety of marker-less motion capture interventions available like the Kinect [10], Kinetisense [32] and human pose estimation (HPE) [52] models. Each of the mentioned methods have their strengths and weaknesses. The Kinect has previously been utilized in a number of VR post-stroke interventions [67]. However, the Kinect is an extra device that must be obtained and set up correctly before it can be used in an exergame [68]. This potential barrier of use can be circumvented by utilizing a method that is solely based on the user possessing a digital camera as found in web cameras, smart phones and tablets [68]. The Kinetisense mainly focuses on analysis of motion and is fairly expensive [52]. Limitations of HPE methods are varying degrees of 3D pose error [58]. HPE provides a method for motion tracking that can be achieved through just the use of a machine learning model and a standard webcam. Motion capture technologies that do not require expensive sensors, but can run on the device of a consumer contribute to a democratization of motion capture technology [23]. HPE methods are computer vision based and do not require specific hardware other than a sufficient computer and a digital camera. Utilizing HPE for motion capture in a system can be practical with regards to scalability, as each user of the system will not need to obtain dedicated motion capture hardware [69]. Several PE tools are available for free, which creates an element of accessibility, as users are not reliant on expensive equipment. HPE technology can be integrated into a game-based system and combined with the use of VR hardware. It is an interesting, novel approach to the task of motion capture in rehabilitative exergame systems.

The not sufficiently met and increasing demand for rehabilitation motivates the exploration of accessible and novel approaches that can assist in rehabilitative treatments for survivors of chronic conditions such as stroke. Virtual reality exergames have the potential to create the appropriate conditions to induce neuroplasticity. Pose estimation can offer accessible and accurate motion capture without demanding the presence of heavy or expensive equipment. Combining these two features in an exergame for stroke survivors can potentially give contribution in the efforts to meet the rising demand for stroke rehabilitation.

1.2 Research question and goal

This project will explore a novel approach to the task of motor rehabilitation for stroke survivors. The project is a proof of concept study exploring the creation of a prototype for an exercise game to train motor function in stroke survivors by combining the use of the two novel technologies human pose estimation and virtual reality. To help guide this process, the following research questions were formulated.

1. How can one proceed methodologically to develop a serious exergame for post-stroke motor rehabilitation?
2. What are the identified challenges and opportunities in utilizing pose estimation for a VR post-stroke motor rehabilitation exergame?
3. What are the key technical considerations required for real-time human pose estimation integration into a VR exergame for stroke rehabilitation?

The goal of this project is to design and develop a proof-of-concept prototype of a virtual reality rehabilitative exercise game that utilizes a 3D HPE model for motion capture. The game will be aimed at stroke survivors with upper-limb motor impairments who are in the sub-acute phase of post-stroke recovery.

1.3 Project structure

This thesis is organized into eight main chapters and structured in the following manner.

1 Introduction

Introduction clarifies the the motivations behind the project, and defines the project's research questions and goals. Lastly, it provides an overview of how this report documenting the project is structured.

2 Background

The background section is intended to provide the reader with information deemed necessary to understand the project's context. The section provides an overview of stroke and motor rehabilitation, and an understanding of the technical concepts exercise games, virtual reality, virtual reality in rehabilitation and human pose estimation.

3 Related work

Related work provides an overview of relevant projects and systems that are somehow topically close to this project. These systems include exercise games, virtual reality systems and rehabilitative exergames. The aim and intention of this section is to provide a clarification of where this project attempts to situate itself in relation to the existing projects in the field, and to help identify what features make this project novel and interesting.

4 Methods

The methods section explains the methodological approach chosen for the project. The section mainly explores the development of the system through the eyes of the selected game design framework, and explores each of the methodological steps taken to advance the project from a conceptual idea to a testable system.

5 Results

Results show the final outcome of the methodological work conducted, from the game design process and the development phase. The resulting rehabilitation exercises selected for the game, the game concept that emerged through the methodological game design steps and the final implementation of the system.

6 Evaluation

An evaluation scheme was deployed in order to be able to provide commentary on the quality of the resulting system. The evaluation section expands on the testing scheme that was selected, the system usability scale (SUS). Further, the section describes the participants of the tests and how the testing procedure was conducted. The section concludes by presenting the qualitative and quantitative results of the evaluation.

7 Discussions

The discussions section offer a reflection regarding the process of the project. The section reflects on the design approach, the use of the PE model BlazePose, limitations of the project, the evaluation scheme and further work.

8 Conclusions

Conclusions answers the research questions presented in the introduction and offers a summary of what was achieved throughout the course of this project and report.

Chapter 2

Background

2.1 Stroke and motor rehabilitation

Stroke is a condition characterised by sudden damage and/ or death of some brain cells, caused by a lack of oxygen supply to the brain [30]. The lack of oxygen can be caused by either inadequate blood flow to part of the brain or spinal cord, or spontaneous hemorrhage to areas of the brain [24, 30]. The outwardly symptoms of a stroke vary greatly depending on the cause and impact of the stroke. Typical symptoms are acute function loss in part(s) of the body such as speech disturbances, weakness of one side of the body, vertigo, headaches and clumsiness. The vast variety of symptoms makes recognizing strokes challenging even for healthcare professionals, which can lead to misdiagnosis [24]. Stroke is a very serious medical condition and can result in death or permanent disability. Worldwide, stroke is a leading cause of both death and acquired disability in adults [30, 42, 14, 16, 24, 31]. Strokes have sudden onsets, meaning the subject of the stroke and their family have no ability to prepare for the consequences [16]. Stroke survivors can experience motoric, cognitive, sensory or social impairments as a consequence of the stroke [36]. Strokes can impair a person's ability to carry out daily activities like walking, driving and writing, thus greatly impacting their ability to be fully independent. It can also greatly impact their ability to participate in social settings and community activities, which can negatively affect their quality of life [36].

Although each case of stroke and its impact is highly individual, some common features are shared by stroke victims. Following the acute onset of a stroke, provided the stroke is not fatal, the stroke victim will enter a rehabilitative phase if needed. The World Health Organization (WHO) defines rehabilitation as "a set of interventions designed to optimize functioning and reduce disability in individuals with health conditions in interaction with their environment" [50]. Rehabilitation is a crucial step for regaining independence and becoming economically active after a stroke [42, 31]. The ultimate goals of rehabilitation are to empower the subject to gain the highest degree of independence in their daily life, and have the ability to partake in meaningful activities such as work, education and family life [50]. A stroke victim can receive rehabilitation during both late and early stages following a stroke. Rehabilitation in both stages have the possibility of providing significant effects for function improvement [14].

Stroke rehabilitation is based on the principle of neuroplasticity, and attempts to encourage changes in the areas of the brain surrounding the affected lesion where the stroke was located [45]. This encourages the understanding that the focus of new improvements in the field of stroke recovery should regard how to optimally engage the structural and functional mechanisms in the brain that facilitate for stroke recovery [45]. Motor function rehabilitation is a subset of stroke rehabilitation that focuses on facilitating for the recovery of motor impairments obtained as a consequence of stroke. Stroke survivors can experience motor impairments that affect their upper-limb function, lower-limb function, mobility, or a combination of these impairments. Between 55% and 75% of stroke survivors experience some degree of motor impairment. Effective motor function rehabilitation requires training that is repetitive, challenging, intensive, task-specific, prominent and motivating in order to best ensure the incurring of neuroplasticity [57].

2.2 Exercise games

Games that require the user to perform bodily movements as key part of the gameplay are called exercise games (exergames) [75]. Oh and Yang [48] analysed the current adoption of the term "exergame" and came up with a more formal definition of the term to better reflects how the word is currently being used. They define an exergame as "a video game that promotes (either via using or requiring) players' physical movements (exertion), that is generally more than sedentary and includes strength, balance, and flexibility activities" [48].

Exergames have several areas of application, and are utilized both for recreational purposes and in the context of serious games. Bellotti et al. [6] defines serious games as "games that are designed to have an impact on the target audience, which is beyond the pure entertainment aspect". This additional aspect the game must inhabit can be anything from the acquisition of new knowledge to the development of new skills [6]. The "serious" objectives of serious games should be stated in the game's design objectives from the very inception of the game's creation [35]. This aspect distinguishes serious games from other types of video games. The market for serious games have grown exponentially in the last decade [35]. Researchers have shown great interest in the application of serious games for a myriad of purposes [18]. Common areas of application for serious games are recreation, education, training, healthcare and more [35, 6]. Regarding the use of serious games in healthcare contexts, the efficacy of serious gaming interventions have been scientifically proven on several occasions. One such example is the use of serious games for treatment of phobias [6].

2.3 Virtual reality

Virtual reality is defined as "the use of interactive simulations created with computer hardware and software to present users with opportunities to engage in environments that appear and feel similar to real world objects and events [71]." Two important concepts relating to VR are immersion and presence.

2.3.1 Immersion and presence

Immersion is a key concept of virtual reality [71]. The concept concerns the degree to which a VR application is successfull in creating a reality that shifts the user's attention from the real world to a new, virtual world [71]. A key characteristic of virtual reality applications is that they attempt to immerse the user in the new reality of the application. Ultimately, the immersion level the user will experience depends on several factors like the quality of the hardware of the VR system being used [71]. Presence refers to the individual's experience of being present in the simulated environment the virtual world provides. Intuitively, presence can be described as the system's ability to facilitate an illusion of "being there" for the user [70]. A high sense of presence has been linked to increased motivation and a more profound emotional response to the simulated environment [71].

2.3.2 Cybersickness

Cybersickness is a somewhat common negative side-effect of the use of VR systems [71]. Weech, Kenny and Barnett-Cowan [70] define the term as "a constellation of symptoms of discomfort and malaise produced by VR exposure." Common symptoms of cybersickness are nausea, headaches, eye-strain, drowsiness and postural instability [71]. Cybersickness is highly individual and can affects users very differently.

2.4 Virtual reality in rehabilitation

Virtual reality-based therapy is therapy that utilizes digital systems to simulate objects and events from the real world [36]. An advantage of VR rehabilitation schemes is that they can provide a platform where patients can practice daily activities that can not be mimicked within a hospital environment [36]. Being able to create simulations of real-world activities is a feature that can activate users' interest and strengthen their adherence to participate in long-term post-stroke treatment [31]. Virtual reality can offer stroke survivors a setting for practicing functional tasks repeatedly in a simulation environment [31]. It has been shown that VR systems can safely be applied in stroke rehabilitation scenarios [36]. Although some VR users experience symptoms of cybersickness like nausea, headaches and motion sickness, these are not factors that should deter the creation of VR rehabilitation schemes, but elements developers and therapists should be aware of and communicate to end-users [36]. The impact of these effects is highly individual and can be dependent on several factors such as the VR hardware, the software, and the user's level of experience with VR [36]. Virtual reality applications can be designed as home interventions, meaning users can engage with them without therapist supervision. This can increase the amount of rehabilitation without increasing the number of healthcare personnel providing the therapy [36]. An increased standardization of assessment and treatment protocols can contribute to ease the impact of the predicted increase in rehabilitation demand [71]. VR exergame systems have the potential to incorporate necessary concepts of neuroplasticity like repetitive, intense and task-oriented training. These features are vital for successfull motor improvement after a stroke [57, 31]. Being able to provide these features in an enriched environment is a great strength of VR interventions for stroke improvement.

2.5 Human pose estimation

Human pose estimation (HPE) is a computer vision based approach to motion tracking that utilizes machine learning (ML) for estimating human movements in an image or video [58]. HPE is an important task withing the field of computer vision, and has several interesting fields of application both within the domains of science and industry. Examples of application fields are human-computer interaction (HCI), human-robot interaction (HRI) and gaming [58]. Human action and motion are two attributes of key practical importance, which has led to continuous focus being put on them and the ability to capture their features from images and video [23]. Recent years has seen the development of a handful of image-based HPE models like OpenPose [27], AlphaPose [17], and BlazePose (Mediapipe) [5, 69]. These models are continuously improving with regards to joint position measurement accuracy [69].

2.5.1 BlazePose

BlazePose is a model for 2D/3D human pose detection and tracking developed by Google Research. Developers of the model, Bazarevsky et al. [5], describe it as "a lightweight convolutional neural network architecture for human pose estimation". During inference, the BlazePose network produces 33 pose landmarks of the subject being tracked [5]. Figure 2.1 shows an image of the 33 pose landmarks the model produces. BlazePose provides accurate and fast 3D human joint position PE [69]. The BlazePose model is part of Google's ML toolkit MediaPipe Solutions, which provides a series of open source libraries and tools for ML projects [69]. MediaPipe is well documented and promotes efficient prototyping for developers [41]. The ML pipeline of BlazePose contains two stages. The first stage of the pipeline is pose detection, and the second is pose tracking [4]. First, the detection stage locates the region of interest (RoI) of the input image where the subject is likely to appear, while the pose tracking step predicts the subject's joints within the RoI [34]. Figure 2.2 shows an overview of BlazePose's inference pipeline.

BlazePose was chosen over competing 3D HPE models due to some inherent strengths that made it a better assumed fit for this specific project compared to alternative options. The model offers more tracking key points than most other pose estimation models on the market [52]. The method extends the Common Objects in Context (COCO) key point topology, which is a commonly used topology for PE schemes. The COCO topology works with 17 key points, which is almost half of what BlazePose works with. The additional key points allow for almost real-time processing of the input image [52], which is significant in a system that depends on real-time response. BlazePose is very suitable for real time motion tracking tasks like exercise applications because it has been specifically tailored for real-time inference, and the network runs at more than 30 frames per second during inference [5]. The BlazePose model outperforms the OpenPose model on fitness use cases [5]. The pose estimation network of BlazePose predicts the location of all the 33 pose landmarks each frame, meaning each pose landmark needs to be visible while the algorithm is operative [5]. This makes it important to place the webcam in such a position that the person being tracked is continuously visible.

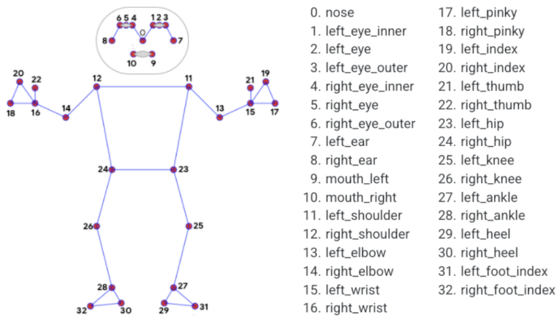


Figure 2.1: The 33 Pose Landmarks of BlazePose [55]

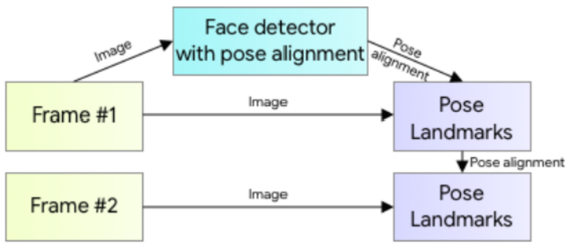


Figure 2.2: Overview of BlazePose's inference pipeline [41]. Frame #1 estimates the subject's RoI. The next frames subsequently utilizes the predictions of the previous frames in order to improve pose tracking [34].

2.6 Development tools

This section elaborates on technical tools that were necessary for the realization of the system. The game engine Unity3D was utilized for VR game development, the head-mounted display HTC Vive created an immersive VR environment, and the webcam LifeCam Studio captured the user's movements on video so it could be given as input to the BlazePose model.

2.6.1 Unity 3D

Unity3D, commonly referred to as Unity, is one of the world's leading game engines. A game engine is a framework used by creators of games and applications. The framework consists of a set of libraries, multimedia management tools and user-interface editors that attempt to ease the task of the game developer [43]. In addition to providing a myriad of practical tools to aid the development process, the framework also implements an abstraction layer between the content of the game and the underlying hardware [43]. Unity consists of a number of different modules and features like its dedicated physics engine and scripting modules. Unity's physics engine can make objects in the game simulate the behavior of real-world objects by following the laws of physics. Scripts are classes that control some behavior in the game [43, 74]. Unity supports the three scripting languages C#, JavaScript and Boo [74]. It also supports development for multiple platforms, a feature partly enabled by the proficient cross-platform performance of all its supported scripting languages [74]. This provides developers the ability to deploy their games and applications on such platforms as Windows, iOS, PlayStation, Android and more [74]. Unity supports VR development and has shown commitment to providing information and resources for VR projects. The Unity manual [66] provides an extensive collection of resources regarding packages, plug-ins and other extended reality (XR) utilities that the engine offers. In addition it is possible to create a new game scene using a VR template that handles much of the initial setup for VR projects. Furthermore, Unity has an active community of both users and employees, and facilitates for a chat forum where developers of different experience level can share knowledge and information [65]. Unity has become an industry leader in the game engine industry with over 47% of game developers utilizing the engine [43].

2.6.2 HTC Vive

HTC Vive is a range of VR headsets and technologies created by the HTC Corporation. The HTC Vive HMD have been shown to enable thorough immersion for users [15]. HTC Vive systems require two base stations also known as "lighthouses" that emit light sweeps [7]. The lighthouses understand where the user is in a space, and alerts the HMD of the physical boundaries of the space [15]. The headset is dependent on the use of long wires that must be connected to both a power outlet and two computer ports. SteamVR is a software tool that enables the use of different VR technologies with the user's hardware of choice. The software supports HTC Vive and was used to setup the HMD.

Figure 2.3: The HTC Vive headset [9].

2.6.3 Webcam

The system required a webcam in order to feed a video stream input to the pose estimation model. A Microsoft LifeCam Studio webcam [11] was utilized for this purpose. Figure 2.4 shows the camera model that was used for the system. The camera has a full 1080p HD widescreen sensor enabling smooth and precise video quality. The wide-angled lens facilitates for capturing a greater degree of the surroundings [11]. Having an external webcam was beneficial as it could be placed in an optimal spot with regards to capturing the whole user on video, which is important to ensure BlazePose can continuously see all 33 joint key points. The LifeCam can be pivoted 360 degrees and placed on a myriad of surfaces, which provided flexibility with regards to the setup environment.

Figure 2.4: The Microsoft LifeCam webcam [11].

2.7 Preliminary study to determine high-level requirements

A preliminary literature study [25] was conducted in preparation of this thesis. The aim of the pre-study was to identify high-level requirements for designing motivational VR post-stroke rehabilitation interventions for motor function recovery. The study concluded in the identification of five high-level requirements, namely meaningful feedback, user safety, user-centric design, rhythmic elements and task specificity. Below follows a brief description of the identified features.

Meaningful feedback refers to the provision of immediate, positive feedback following a performed task. User safety is a should always be addressed. The dynamic motion aspect of exercise games pose a risk for users wearing HMDs for immersive VR. This risk can be diminished by concentrating the exercise intervention on upper-limb motor rehabilitation, as the lower body can then be somewhat stationary during exercise performance, either by taking a sedentary position or by standing still. To mitigate potential cybersickness, users should be advised to not use the headset continuously for too long at a time, but be mindful to take regular breaks. User-centric design refer to implementing variables that make the gaming intervention personalized to the user. The task difficulty of the performed exercise is an important factor to personalize. The user must experience the practiced task as challenging in order to engage in active problem solving and induce new learning. Increasing difficulty to follow the user's progression has shown positive results in post-stroke motor recovery. Rhythmic elements are a more novel concept that can be incorporated in motor function rehabilitation. The potential effect of music depend on the diagnosis of the individual subject, but music generally has positive effect with regard to cognitive outcomes. Rhythmic cueing can have positive effects for people in stroke recovery, and can deter tendencies of user passivity. Task specificity refers to rehabilitation exercises being based on tasks matching activities found in the user's daily life, known as activities of daily living (ADL). Task-specific practice is a critical principle in neurological rehabilitation. In addition to basing exercises on ADL, the exercises should be executed repetitively, and challenging to the user.

Chapter 3

Related Work

The following section presents projects that are topically similar or adjacent to the work documented in this report. The section intends to provide some insight into how this project relates to work that has already been conducted in this field of research.

There exists several examples of post-stroke rehabilitation exergame interventions that utilize virtual reality technologies. Motion Rehab AVE 3D [67] is a VR exergame for post-stroke rehabilitation. The game focuses on upper-limb motor rehabilitation and balance rehabilitation. The system depends on the Kinect for motion capture and the intervention implements in total six training exercises. Participants can engage with the system by either using a TV screen, or enable immersive VR through the use of a HMD. Subramanian et al. [62] developed a VR system for arm motor recovery, also intended for chronic stroke patients. The system simulates a supermarket setting and participants practice exercises by targeting groceries and reaching for them repeatedly. The training exercises were conducted from a sedentary position. Ferreira and Menezes [18] also implemented a VR serious game targeting post-stroke upper-limb motor rehabilitation. The game was evaluated by therapists and went through a pilot study with patients. The study resulted in favourable results, with the system being very well received by the patients who tested it. The system was deemed accurate for training about 75% of the clinical cases treated at the rehabilitation clinic where the tests were conducted.

There are several examples of HPE models being implemented into working prototypes for various purposes. Wang et al. [69] designed an exergame for the rehabilitation of knee disorders that utilizes the HPE model BlazePose for motion capture. The system was evaluated to be both a positive user experience for patients, and an efficient rehabilitation tool for knee disorders. The pose estimation aspect of the system provided accurate and smooth interaction and motion recognition [69]. Several systems have been developed for the use of HPE in the fields of sport and physical exercise [1]. Badiola-Bengoa and Mendez-Zorrilla [1] conducted a review of camera-based HPE applications in these fields. They were able to identify 20 relevant papers that fit their inclusion criteria. The papers all centered sport and physical exercise as application areas, but the specific intentions behind each system varied. Some noticeable examples of interesting systems were the use of 3D HPE for estimating the gait parameters of elderly users for early detection of dementia [33]. Another interesting system was that of Wu and Koike [73] who used 3D HPE with a webcam for mixed reality (MR) martial arts training. These systems both show HPE being utilized for real-time exercise applications. The latter also included the use of a HMD to enable a MR environment for the user. Another HPE intervention targeting elderly users is the NeuroPose [56]. The system aims to make preventative care more accessible for the elderly by providing web-based exercise routines where the user executes training exercises and receives performance feedback. This system also aims to incorporate a social aspect where users can connect in an online community. PoseTrainer [8] is a system for tracking and detecting a user's exercise pose to provide detailed feedback in the form of corrections and recommendations for how they can improve their form. The system utilizes HPE and is currently implemented for four training exercises, with opportunities for future extensions. Pauzi et al. [52] implemented a system for detecting movements that can potentially give rise to physical injury over time. The system calculates several metrics relevant for determining if the execution of the movement can result in injury over either a short or long-term period.

There has been considerable activity on the topic of VR interventions for stroke rehabilitation. The identified opportunities linked to the use of VR in the context of stroke recovery has led to much interest in exploring this path of research, and resulted in the development of several interventions for this purpose. There has been a steady increase in research into the topic of HPE and its potential areas of application [76]. Pose estimation is as of yet not very widely applied in the field of rehabilitation, but there are examples of HPE systems for both physical exercise purposes and elderly users. A very attractive aspect of enabling HPE in rehabilitation systems is that the potential hardware requirements are very low, for example only requiring the user to own a webcam. This opens the door for the creation of accessible and scalable interventions. Employing motion capture in the form of 3D HPE in the context of motor rehabilitation for stroke survivors is a novel path of exploration for the creation of future post-stroke interventions. Combining these two technologies creates the foundation for an intriguing path of research that has not been significantly explored previously.

Chapter 4

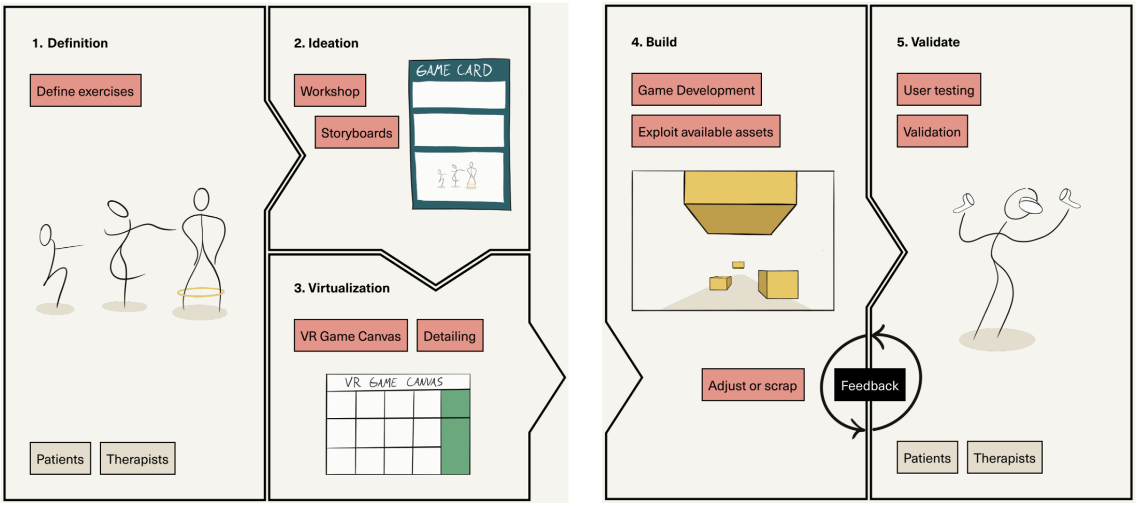
Methods

4.1 Game design process

The design of the game was created through a game design process based on the design framework of Støen and Fridheim [59]. The framework is specifically developed for creating serious games for rehabilitation. The framework includes explicit steps for incorporating XR elements into he game. The game design framework consists of 5 main stages. The stages are definition, ideation, virtualization, build and validation. The stages of the framework are meant to be executed sequentially, with the results of each validation stage serving as feedback for the next build iteration. The framework and its main stages are visualized in figure 4.1.

The framework of Støen and Fridheim [59] builds on the methodology of Pirovano et al. [53] for designing safe, therapeutically valid exercise games. A key feature reflected in both frameworks is the clear line drawn between exercises and games. Pirovano et al. [53] creates a definition for exergames that highlight the relationship between an exercise and a game. They define that "an exergame is an exercise with a game built into its structure", and state that "only if all the aspects of an exercise are fulfilled can a therapeutic exergame be considered really effective and safe for the patient." In this manner, the methodology puts equal focus on the therapeutic outcome of a rehabilitation game as on the game itself.

The framework was a suitable choice for this development process as it is specifically tailored for the development of rehabilitative exercise games and includes considerations for incorporating XR elements. The framework promotes fast development and implementation, and the creation of systems that encourage improved rehabilitation [59]. This further coincided with the time scope and overall intentions of this project. The five main stages of the framework can be adjusted both in length and scope to fit the desires and requirements of individual projects. This also includes flexibility regarding the degree of stakeholder involvement. This is a very practical consideration as stakeholder involvement heavily depends on stakeholder availability, and getting therapists and patients involved can be a complicated process. So although including patients and therapists in stages of the framework is encouraged, the execution of the design process can proceed without them. As this system is still in a preliminary stage, neither stroke survivors or physical therapists were contacted, as the focus was mainly on the proof-of-concept aspect of combining novel technologies. This made the adaptability of the framework very beneficial, as all stages could be completed without these stakeholders.



(a) (b)

Figure 4.1: Framework of Støen and Fridheim [59] for developing serious games for rehabilitation.

4.2 Exercise definition

The first stage defines the exercises the game will be built around. Several considerations needed to be made regarding the choice of exercises. To be considerate of user's safety, upper-limb exercises were mainly considered for integration into the game. Exercises focused on the user moving around in space were considered especially unsuitable, as it could be unsafe for the user to execute without supervision, and also impractical with regards to the XR aspect of the system. Based on these considerations the choice of training exercises was narrowed to only include upper-limb exercises.

The Stroke Rehabilitation Assessment of Movement (STREAM) [28] is a tool for evaluating the performance improvement of user's who undergo motor rehabilitation following a stroke. The tool presents an overview of current movement exercises for users in post-stroke rehabilitation selected by a panel of therapists. The exercises are grouped based on the sub-group of limbs they train, and sorted within their sub-group based on ascending difficulty logit measure [28]. The standard upper-limb exercises from the STREAM are shown in figure 4.2. Two exercises from the upper-limb movement sub-group of the STREAM were chosen to form the therapeutic basis for the exergame. The exercises are therapeutically valid as they have been chosen by physical therapists for the specific purpose of stroke motor rehabilitation. The first exercise chosen was "Raising arm to fullest elevation", which has a difficulty logit measure of 1.50. The second exercise was "Raising hand to touch top of head" and has a difficulty logit measure of -0.46. Stroke motor rehabilitation exercises should map to the execution of daily tasks, which was the case for both the chosen training exercises. "Raising arm to fullest elevation" maps to daily activities like waving, cheering, reaching for items on a shelf and drawing blinds. "Raising hand to touch top of head" maps to the daily activities putting on a hat, checking ones temperature, wiping sweat and applying shampoo.

The two chosen training exercises considered the user's safety, were therapeutically validated and mapped to ADL, making them a stable choice for motor rehabilitation.

Item[a]	Difficulty Logit	SE
Upper-limb movement		
Elbow extension while supine	−0.77	0.30
Raising hand to touch top of head	−0.46	0.30
Scapular protraction	−0.02	0.31
Making a fist	1.21	0.32
Moving hand to sacrum while sitting	1.29	0.32
Raising arm to fullest elevation	1.50	0.32
Supination and pronation	1.78	0.33
Total extension of fingers	2.06	0.33
Threshold 1	−2.29	0.10
Threshold 2	2.29	0.10

Figure 4.2: A selection of standard upper-limb exercises for motor rehabilitation of stroke survivors taken from the STREAM evaluation tool. Exercises are shown with their difficulty logit measures and the associated standard error (SE) Hsueh et al. [28].

4.3 Concept ideation

The ideation stage consists of a formalized brainstorming session that utilizes the tools workshop and storyboarding to generate ideas for the game concept [59]. A digital workshop was carried out with two participants. The workshop was based on the two upper-limb exercises chosen from the exercise definition stage. Participants used miro-board templates made available from [59] to express and visualize their ideas. The outcome of the workshop was three separate game concepts. The first concept was a maestro/ orchestra conductor game where the user would hit notes above their head to the beat of the music. The second concept was a fruit-catching game where fruit would fall from the sky and be caught by the user, who would then deposit it in a basket on their head. The third concept was a volleyball game, where the player would attempt to keep a volleyball afloat by continuously serving it. All concepts incorporated both of the chosen rehabilitation exercises.

The volleyball concept was the first concept to be excluded from further exploration. This was due to the concern that the consolidation between the second rehabilitation exercises "raising hand to touch top of head" was not ideally combined with the abstraction of a volleyball game, and might turn out to be somewhat unnatural. The maestro concept was not chosen due to a concern that it would be difficult to find suitable freely available assets for the game aesthetic. This concern was brought about due to a preliminary search in the Unity asset store [64], where the search words "maestro" and "orchestra" did not result in any suitable assets in the 3D-objects category. The fruit catcher game concept was ultimately selected as the most suitable for further development. Several suitable packages were found in the Unity asset store that matched the intended visual look of the concept, and the concept worked well with the rehabilitation exercises.

4.4 Virtualization

Pirovano et al. [53] describes virtualization in the context of therapeutic game design as the task of implementing the goals of a physical exercise into a virtual exercise by defining input and output requirements. The virtualization stage analyses the chosen game concept from the concept ideation stage through a game canvas specifically developed for serious games that incorporate VR elements. The blank VR game canvas is shown in appendix C. The canvas includes fields to clarify how the game concept will work in terms of game mechanics, and how crucial concerns such as safety will be maintained in the game. First, the canvas defines fields specifying the user's inputs and outputs in terms of in-game actions. The canvas also makes special consideration to capture details regarding the game concept, and fields specifying requirements that pertain to serious games, which is the nature of games the canvas is intended for [59].

4.5 Build

The game was developed with the Unity3D game engine. The game concept fruit catcher was implemented into a working prototype through the following build process. All referenced scripts can be found on the project's github [26].

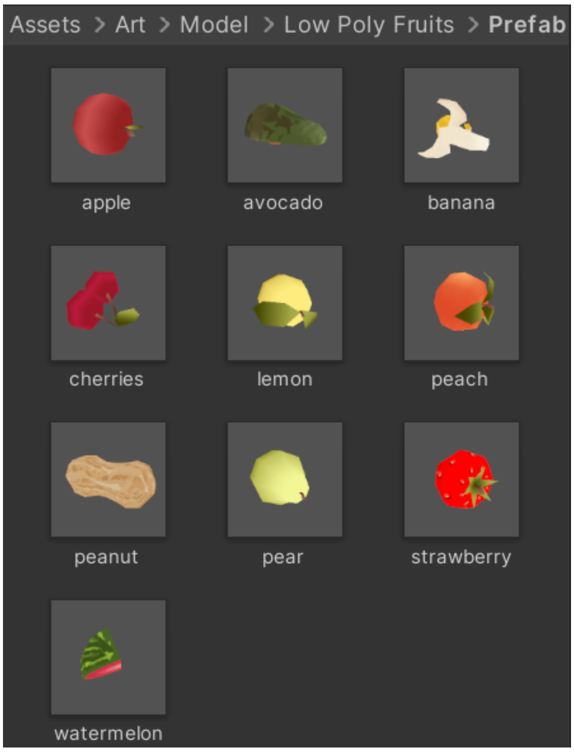


Figure 4.3: Low-poly fruits GameObject Prefabs

4.5.1 Gameplay

The gameplay of the exergame is centered around the catching of falling fruit objects. 3D fruit-objects from the package *Low Poly Fruit Pickups* [60] were set to serve as the falling fruit components. The package contains ten 3D fruit elements, all of which were utilized in the game. The fruit objects come as Prefabs which are template objects that can be stored with all properties already decided [66]. This enabled the fruit prefabs to be used as templates for every fruit object that was instantiated during the gameplay. The fruit asset elements from the package are shown in figure 4.3. A spawn script was implemented to randomly spawn the different fruit objects in a set radius around the player. The fruits were set to spawn at random locations within the given radius, with equal probability of spawning on either side of the player. This was to ensure the game was not biased towards training one of the hands more than the other, but providing both arms could get equal amounts of training. The wait time between each spawned fruit and the total amount of fruits that spawn in a session can be adjusted in the difficulty settings interface, controlled by the script DifficultySettingsManager. The fruit objects have three scripts attached to establish desired behavior. The first script, StopOnCollision, ensures the falling fruits stop when they are caught by one of the player's hands. Catching a fruit adds a point to the scoreboard for the first exercise. The script Feedback manages the visual aspect of giving the user feedback. The second script, FollowPlayer, makes caught fruit follow the hand that captured it. The third script, PlaceInBasket, deposits the fruit in the head basket when the player places it there, and awards the player a point for completing the second rehabilitation exercise "raising hand to touch top of head". For when the player is not able to catch the falling fruit, the script DestroySpawnedObject destroys fruit objects that have not been caught as they ultimately collide with the ground. In addition to the scripts, the fruit objects have an audio source attached that plays an audio clip each time a fruit is spawned.

4.5.2 Environment

The most fundamental elements of the environment were created using a simple Unity skybox that was assigned a blue background, and a basic plane element. The scene composition comprised of nature elements that fit the theme intended for the environment of the game concept, fruit catcher.

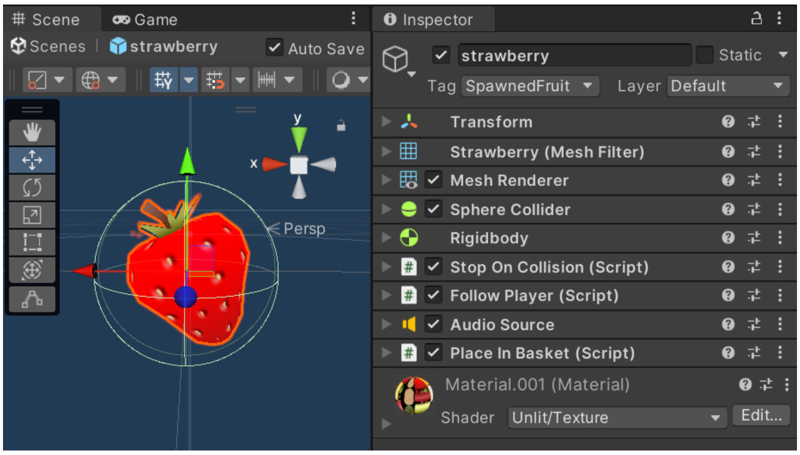
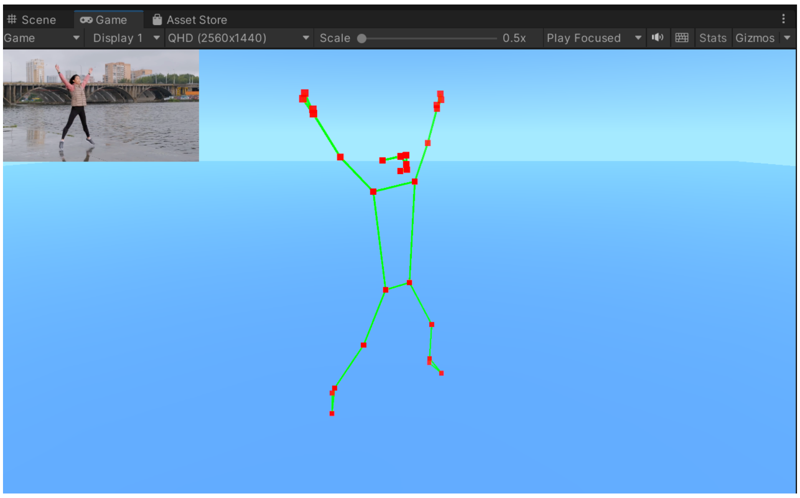


Figure 4.4: Scene- and inspector view of the strawberry prefab

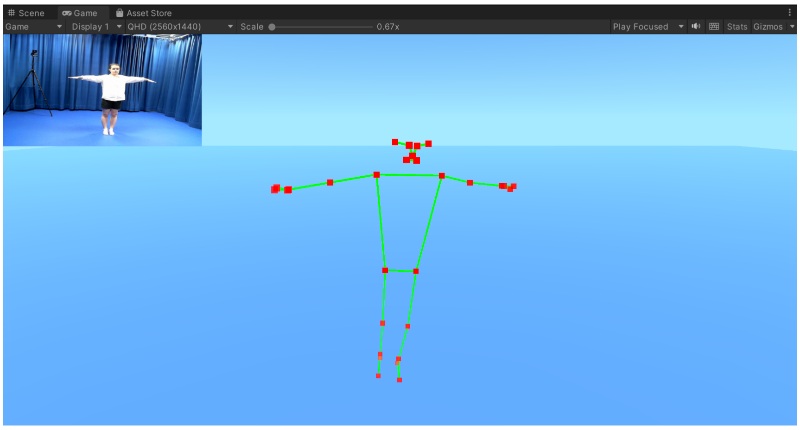
Freely available 3D-object assets were acquired from the Unity asset store [64]. These contributed in creating the desired visual look for the game. Objects from the package *Free Low Poly Nature Forest* [54] were placed around the plane to create a natural environment. To give the environment a positive, uplifting mood, the chosen nature elements had a bright, colorful appearance. The use of low-poly 3D objects set the tone for the visual "feel" of the game environment. 3D objects with low-poly meshes generally assume a visual look that is further from "real world" objects than 3D objects with high-poly meshes. To create a consistent visual atmosphere, the use of low-poly 3D objects was carried through to the other 3D objects that were used to create the environment, as well as the avatar.

4.5.3 BlazePose

A method for full body tracking was implemented by integrating Mediapipe's BlazePose model into Unity. This was achieved by utilizing Unity's own ML engine, Barracuda. The Unity manual describes the Barracuda package as a "lightweight cross-platform neural network inference library for Unity [63]." The Unity package *BlazePoseBarracuda* developed by github user creativeIKEP [13], and made available through an apache license [20] runs the BlazePose pipeline on the Unity Barracuda. The input of BlazePoseBarracuda can be set to either an image, a recorded video or a live video stream from a digital camera. For this system it was set to be the live video stream from a high definition web camera. The web camera was placed in a position where it was assumed it could sufficiently capture every pose landmark on the user at all times. Barracuda and BlazePoseBarracuda were both installed through the Unity Package Manager. Figure 4.5 shows BlazePose running in Unity on two different inputs, one the recorded video of a woman jumping continuously, and the other a live webcam stream of a person standing in T-pose. BlazePose detects the pose landmarks of the person in the recorded video and on the webcam, and draws a simple stick figure representing the person based on the landmark key points.



(a) BlazePose running in Unity on a recorded video input of a person jumping



(b) BlazePose running in Unity on a live webcam input of a person in T-pose

Figure 4.5: BlazePose running in Unity on two different input sources.

4.5.4 Humanoid avatar

BlazePose provides a simple line figure created from drawing a green line between each of the pose landmark key points the model tracks, as seen in figure 4.5 and ??. The choice was made to implement an avatar with more humanoid features to create a stronger sense of association between the user and the figure. The avatar was required to be neutral in terms of physical traits to avoid users experiencing diminished attachment if their individual physical traits did not match the physical traits of the avatar. A human skeleton was chosen to fit the criteria of being in the intersection between humanoid and neutral in terms of physical features, and was thus chosen as the avatar. The 3D model for the skeleton was acquired from the package *Lowpoly Medieval Skeleton* [61]. Incorporating an avatar with a low-poly model was also beneficial with regards to computer performance, and matched the visual atmosphere of the game environment.

Controlling the avatar was achieved by building a solution for mapping the joints coordinates maps detected by the BlazePose model with the positions of the avatar's joints. This solution was implemented in the script VNectModel. The first step of the solution is to initialize the orientation of the joints of the 3D avatar model. This must be done to compensate for the fact that the local orientation of the skeleton model's joints are not consistent with the orientation of the global world space of the Unity scene. Inconsistency between the local orientation to joints of 3D humanoid models in Unity, and world space orientations is a common occurrence [44]. Creating a compensation for this is important to avoid joints pointing in directions that are neither consistent with the world space or with each other [44]. Some of the pose landmarks BlazePose detects are situated in the face of the user. The use of a HMD hinders these landmarks from being detected

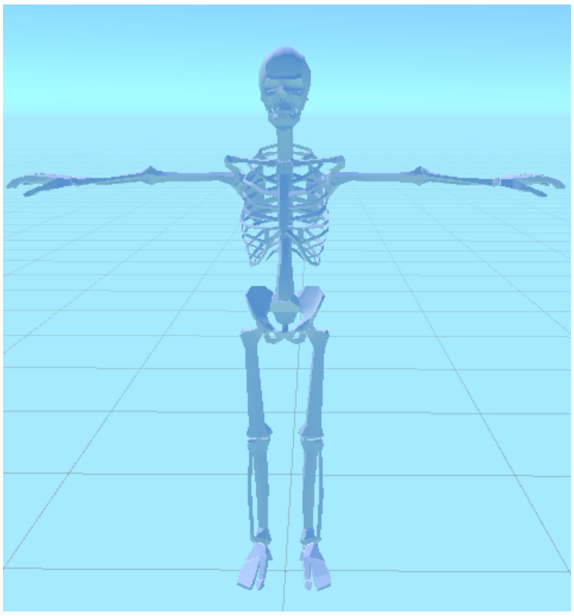


Figure 4.6: Scene view of the low-poly skeleton avatar in T-pose

accurately. This poses a challenge for the estimation algorithm, and can result in glitching. To overcome this limitation, the head position of the avatar is instead found using the position tracked from the HMD.

Miller [44] conducted an analysis of methods for utilizing 3D pose estimation models with humanoid avatars. One of the analysed methods was the solution *ThreeDPoseUnityBarracuda* created by Digital Standard Co. LTD. [40]. The solution creates a standard method for using 3D humanoid avatars with 3D HPE models and is licensed for non-commercial use. It also takes into account the aforementioned challenge regarding joint orientation inconsistency. Odrowaz-Sypniewski and Lamda [47] created the project *FullBodyPoseEstimation* with the intention of creating a solution for full body pose estimation for VR applications made available under a General Public License [19]. Part of the solution is creating a 3D avatar to be used in VR with the BlazePose model. The solution of which is based on the *ThreeDPoseUnityBarracuda* project of Digital Standard Co. LTD. [40]. Elements from both these projects were takes as inspiration for creating the solution for the final skeleton avatar. Figure 4.6 shows the final skeleton avatar in T-pose.

4.5.5 Virtual reality

An immersive virtual reality experience was realized with the use of a HTC Vive HMD. SteamVR [12] set up the connection between the computer and the HMD/ lighthouses. Through the Unity tool XR Plug-in Management, the XR plug-in provider OpenXR [29] was installed into the project. OpenXR is an open source standard for cross-platform XR development and works with the HTC Vive technologies.

4.5.6 Setup

The setup for the technical implementation consisted of one high definition web camera, one pair HTC Vive headset and one pair of HTC lighthouse base stations. The web camera and HMD were connected to a computer with the softwares Unity and SteamVR running. The final implementation of the game and BlazePose runs in Unity.

4.6 Validate

User-tests were conducted with eleven test participants over the course of two days. Feedback was gathered from the participants through a usability schema and general feedback. The validation step is discussed in section 6: Evaluation.

Chapter 5

Results

The work conducted in this thesis produced a game concept for a post-stroke VR rehabilitation exergame, a technical setup for the exergame and a prototype of an implemented system.

5.1 Game concept

The result of following the methodological design and development process was a concept and a prototype of a VR exergame for upper-limb stroke rehabilitation. The game concept is summarised through the following sentence "A fruit-catcher VR game where the player gathers fruits and places them in a fruit-basket on their head to gather as much fruit as possible." The game is based around the training exercises "Raising arm to fullest elevation" and "Raising hand to touch top of head". The motions of the user are captured using the human pose estimation model BlazePose. The fruit catcher exergame concept is based around two aforementioned exercise activities. The user is situated in a natural environment, inspired by forest/ garden elements like trees, flowers and mushrooms. The user engages in two main activities. First, falling fruit is caught by stretching arms upwards. Secondly, caught fruit is placed in a basket on the user's head. Through these activities the concept incorporates the training exercises. The challenge for the user is to catch every piece of fruit and collect them in their basket, this is also their main conceptual goal. The player is the only character in the game concept and the game is single-player. The user is immersed in virtual reality through the use of a HMD and is represented in the game through an avatar. The game has a cheerful atmosphere that is neither particularly relaxed nor tense, providing a neutral environment. The design elements are semi-realistic and somewhat humorous. All the described elements of the generated fruit concept are visualized in figure 5.1.

The results of the virtualization analysis of the fruit catcher concept mapped features of the concept to more concrete game mechanics. The game input is based on the BlazePose tracking of the user executing the training exercises. The output of completing the exercises is feedback in the form of points. The anticipated utilitarian learning outcomes of the game is for the user to regain motor function, and contributing to the user achieving the highest possible level of independence in their daily life. Additionally, the game aims to empower those undergoing post-stroke rehabilitation and provide them with a more engaging exercise environment for training repetitive, task-specific exercises. The results from the VR game canvas analysis of the fruit garden concept is further shown in figure 5.2.

5.2 Exergame setup

The resulting technical setup is shown in figure 5.3. It consists of two HTC Vive base stations placed diagonally on each side of the player, a HTC Vive headset, a LifeCam Studio webcam and a computer running SteamVR and the game on Unity. The headset is plugged into the computer through a display port and a USB port, and connected to a power outlet. The camera is connected to the computer through a USB port.The base stations are both connected to power outlets. The HMD and the base stations connect to the computer's software with SteamVR.

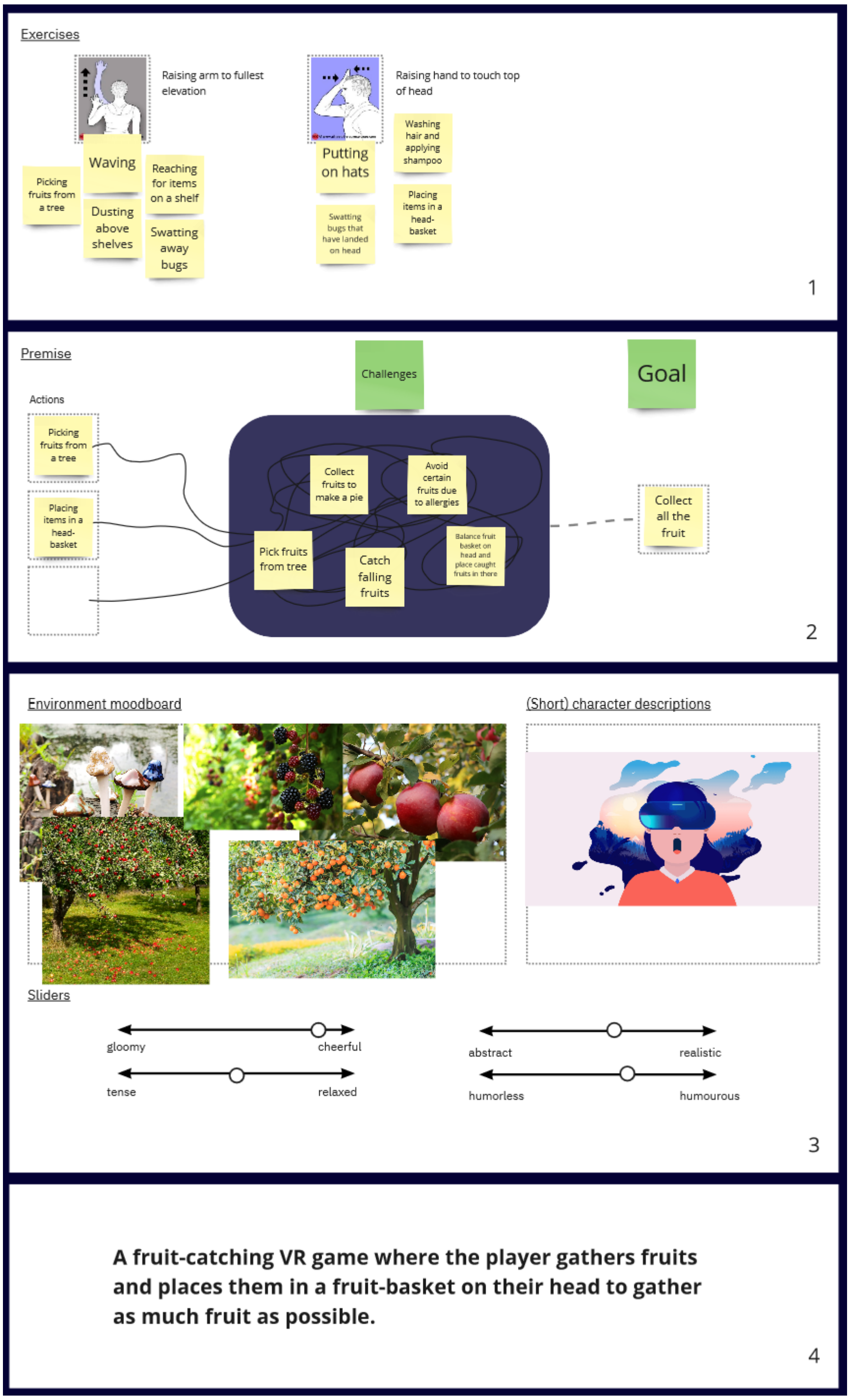


Figure 5.1: Fruit catcher game concept generated through the Støen and Fridheim [59] framework. All images accessed through creative commons license.

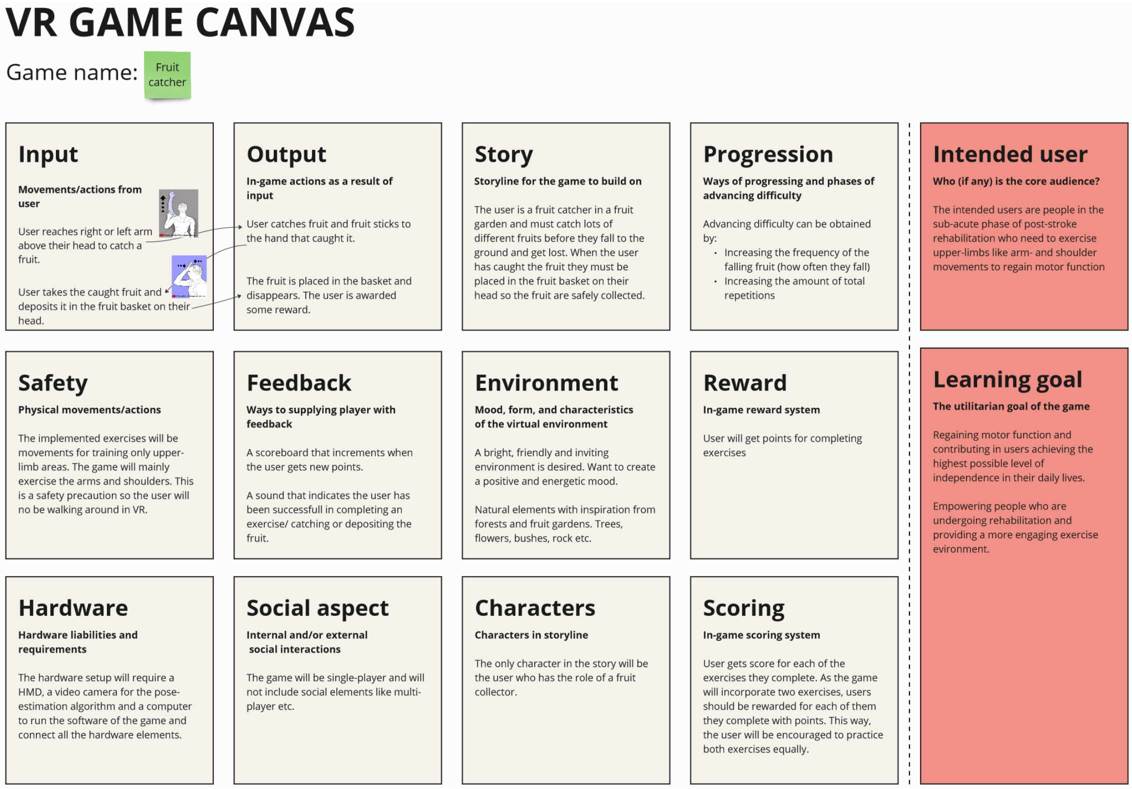


Figure 5.2: Result of the virtualization analysis of fruit catcher Støen and Fridheim [59].

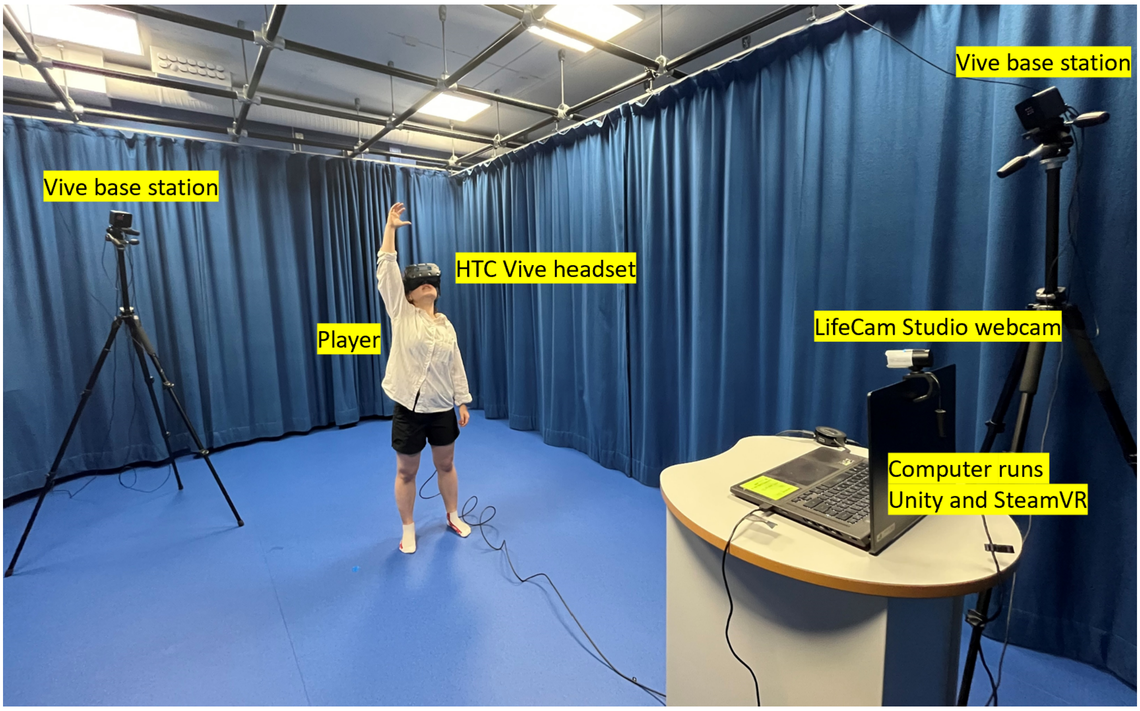


Figure 5.3: The final technical setup for the exergame.

5.3 System implementation prototype

Figure 5.4 shows the view of the prototype from both the hierarchy and the scene view provided by Unity. Figure 5.5 shows the prototype only from the scene view of the game. The figure shows player being centered in the plane, and represented through the humanoid skeleton avatar. The fruit collection basket is a permanent fixture on the head of the avatar.

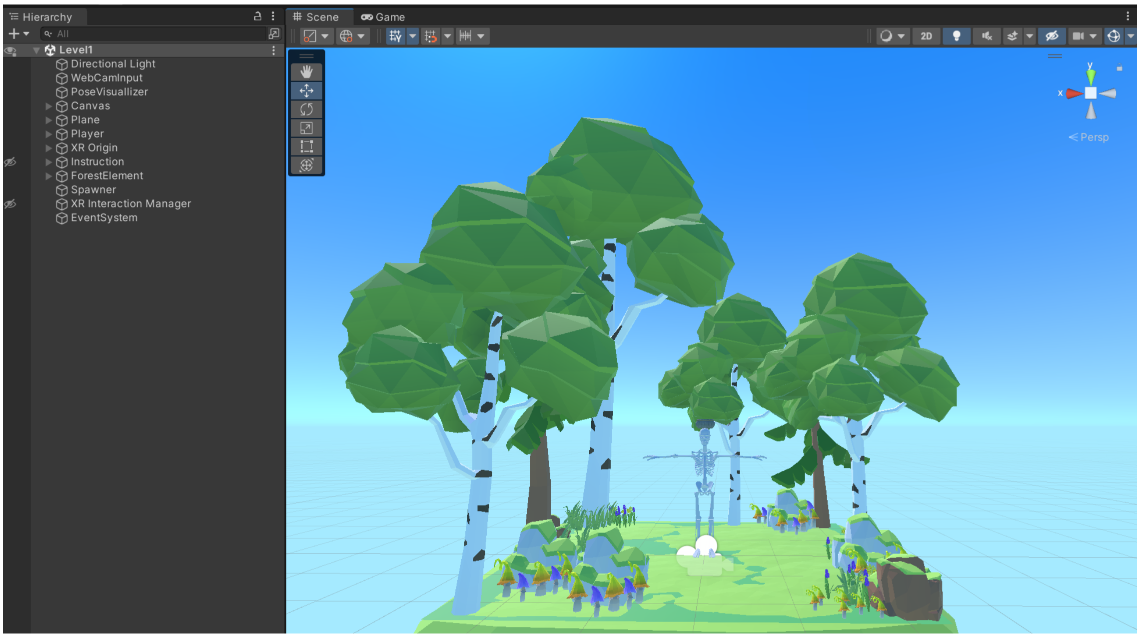


Figure 5.4: Forward position of the player: Hierarchy and scene view

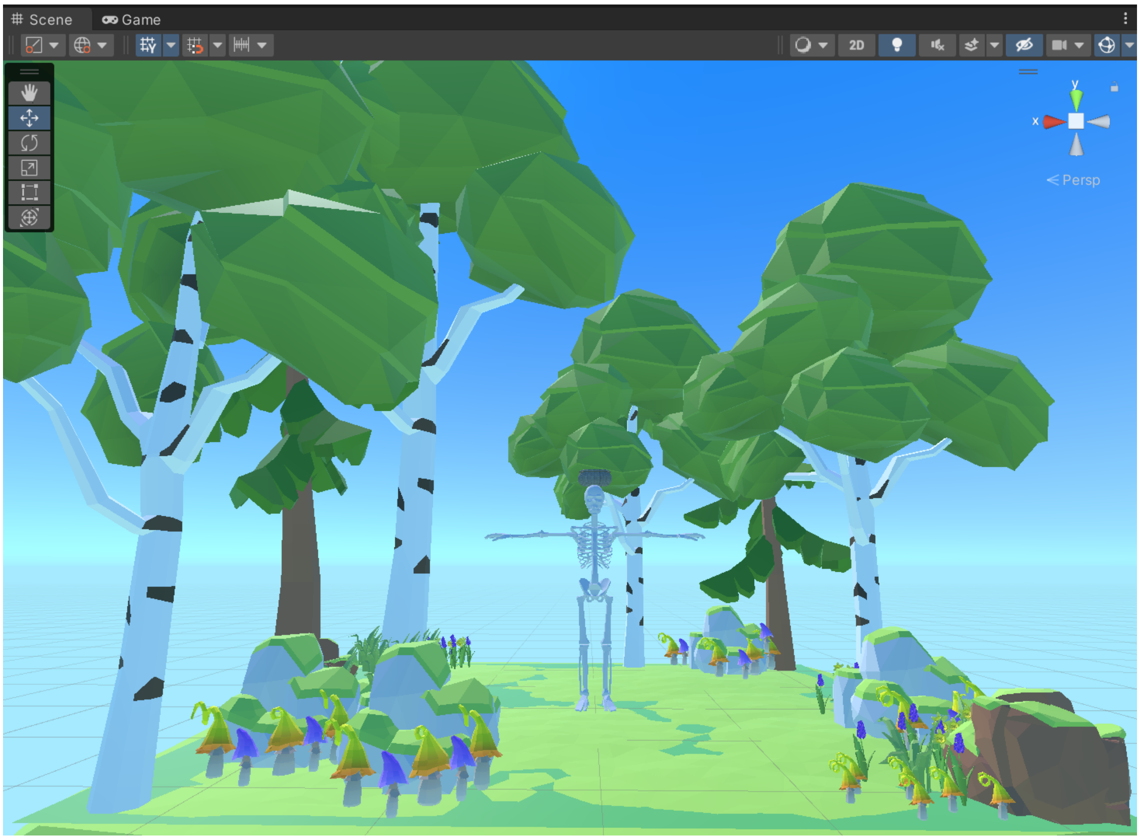


Figure 5.5: Forward position of the player: Scene view

Chapter 6

Evaluation

The objective of the evaluation stage was to conduct a preliminary evaluation to understand the quality of the system in terms of its technical viability. Measuring the usability of the system using a standardized usability test was deemed a suitable approach for achieving this. Assessing usability is an important part of the development process when designing new virtual reality systems with therapeutic intentions [36]. The evaluation phase resulted in both quantitative and qualitative results. The quantitative results were the the scores from the usability schema. The qualitative results were written feedback in the "additional thought/ comments" section of the usability schema, oral feedback and observed actions.

6.1 The System Usability Scale

The System Usability Scale (SUS) [B] is a survey scale for measuring the perceived usability of a system [3]. Perceived usability is a key metric for evaluating user experience [38]. There are several qualities of the SUS that make it a solid choice for measuring general usability. The SUS is flexible enough to be applicable for measuring the usability of a great variety of technological systems. The questionnaire can be answered relatively quickly, and the final result is a single score, making the result easily interpretable [3, 2]. Additionally, the SUS has since its introduction in 1996 been well tested and is proven to be an accurate tool for measuring perceived usability. Research has been conducted to explore the reliability, sensitivity and validity of the SUS, with steadily favourable results [38]. Today, the SUS is the most widely used tool for measuring perceived usability, and can be considered an industry standard [3, 38]. The aforementioned strengths combined with the time frame of the design, implementation and testing of this system made the SUS a suitable choice for evaluating the system's initial usability.

The SUS consists of ten statements regarding the system being tested. For each statement, the user provides a score regarding the degree to which they comply with the statement. Scores are given on a 5-point scale, where 1 represents the lowest strength of agreement (strongly disagree), and 5 represents the highest strength of agreement (strongly agree) [3]. The final individual score results of the SUS are interpreted using a score range of 0 to 100, where higher scores imply better usability. This means each question carries a weight of 10 points. Each item on the SUS needs to have been evaluated for the final score to be valid [3, 38]. To ensure this during user testing, each item was made to be obligatory for the user to fill out when answering the questionnaire.

The statements of the SUS alternate in tone. Each of the statements with an odd index are positively-toned, while the statements with an even index are negatively toned. To account for this in the calculation of the final score, the following adjustments are made. The odd-indexed statements' contribution to the user's final SUS score is found by subtracting 1 from the score the user gave to each specific odd-indexed statement. The even-indexed statements' contribution to the user's final score is found by subtracting the score the user gave a particular even-indexed statement from 5. Undergoing this conversion means that for each statement the minimum value becomes 0 and the maximum value becomes 4. After converting the score range of the odd-indexed and even-indexed statements as explained above, the user's SUS score is found by taking the sum of the converted scores from each of the ten statements. This creates a score scale that ranges from 0-40 for the user's individual SUS score. Because the score scale should ultimately range from

0-100, the user's final score is multiplied by 2,5. Ultimately, the final SUS score of the system being tested is found by taking the average score of the final individual SUS scores of each participant [38]. Interpreting a SUS score in terms of adjectives is done in the following manner. A SUS score of minimum 70 can be interpreted to indicate good usability [2]. Scores below 50 are deemed unacceptable in terms of usability, while systems that score above 90 are considered highly superior [3].

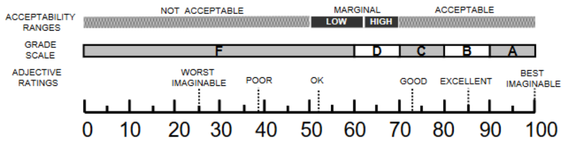


Figure 6.1: Adjective rating scale for SUS scores [2]

When interpreting results of the SUS, one must keep in mind that the score the user assigns each statement is essentially half of the given value that the answer reflects in the final score. Thus, one should not conclude that a system with a score of 50 is "half as good" as a system with a score of 100. But rather that the SUS has demonstrated the former to contain some essential usability flaw [3]. Additionally, the SUS can more easily demonstrate that a system is unacceptable in terms of usability than demonstrate that a system is acceptable. If a system scores below 50 on the SUS, one can claim with a high degree of confidence that the system has some crucial inadequacy regarding the quality of usability. However, if a system scores above 70, one can not guarantee with the same degree of confidence that the system is highly acceptable in terms of usability [3]. Research has deemed the SUS to be seemingly gender neutral, but a slight significance has been shown between a user's SUS score and their age. The results identified that higher age has slight correlation with a more negative SUS score [3].

6.2 Participants

Eleven participants were recruited for participation in the user tests of the final implementation. The participants consisted of people recruited from my network of acquaintances. Participants were provided with a consent form A.2 to explicitly state their desire to participate in the testing. All of the eleven invited participants complied with the statement of consent on the consent form. The form collected some information pertaining to the demographic of the participants. All participants were born between 1998-1992, and consequently reside in the age range 24-31. The represented age span is thus not particularly wide, but this was not deemed problematic as the testing is still in a preliminary stage. Out of the three options male, female or other(s), 4 participants stated their identified gender to be male, 7 stated they identified as female and 0 identified as other(s). All participants replied negatively when asked if they had any serious motor function disabilities.

6.3 Testing procedure

Usability test were conducted with the intention of testing the technical viability of the system. I was especially interested in getting feedback on how the integration of the different technical elements like the pose estimation and the HMD worked together from an outside perspective. It was decided to not attempt to recruit test users with motor function disabilities, even though they are the intended user group. The reason for this was that the system is still in an early phase of implementation, making it more sensible to focus on getting feedback regarding its general usability. Another factor was that getting approval from the Norwegian Center for Research-data (NSD) to conduct user tests with stroke survivors with motor function disabilities would be significantly more time-consuming. This was an important factor considering the time-scope of this thesis and the time that was set off for completing the user tests.

User-testing was conducted in the following manner. The testing consisted of three main stages. First, upon entering the lab, participants were invited to read and sign an information- and consent form A.2. Only if the participant chose to comply with the consent form did the test advance to the next stage. The intention behind the form was to confirm the participants interest in being a part of the testing, and making sure they were aware of their rights. An application was submitted to NSD in December 2022 requesting permission to conduct usability tests of the system using the SUS questionnaire. The application was approved prior to the execution of the user tests. The second stage was actually playing the game. Participant's were provided some practical information regarding the use of the HMD, and informed that the camera would be tracking their movements for input to the BlazePose model. Participants played the game for 10-20 minutes at varied difficulty level. As all participants were healthy, the easier levels were a bit too easy for most of them. Finally, after testing the system, participants answered the SUS questionnaire anonymously.

6.4 Evaluation results

6.4.1 Quantitative results

Figure 6.2 shows the average score participants gave to each statement of the SUS with the belonging standard deviation. The figure shows the original score participant's gave (1-5), before any conversions were made. Statement 6) "I thought there was too much inconsistency in this system" had the largest standard deviation, meaning it was the statement to which the participants gave the most varied range of answers. Statements 2) "I found the system unnecessarily complex" and 10) "I needed to learn a lot of things before I could get going with this system", both had a standard deviation of 0. These were the statements where users were most in agreement. Figure 6.3 shows the final individual SUS score for each of the eleven test participants. Based on the eleven test participants, the final SUS score of the system was calculated to be 83,2. This puts the system in an acceptable range in terms of usability, based on the guidelines for interpreting SUS scores.

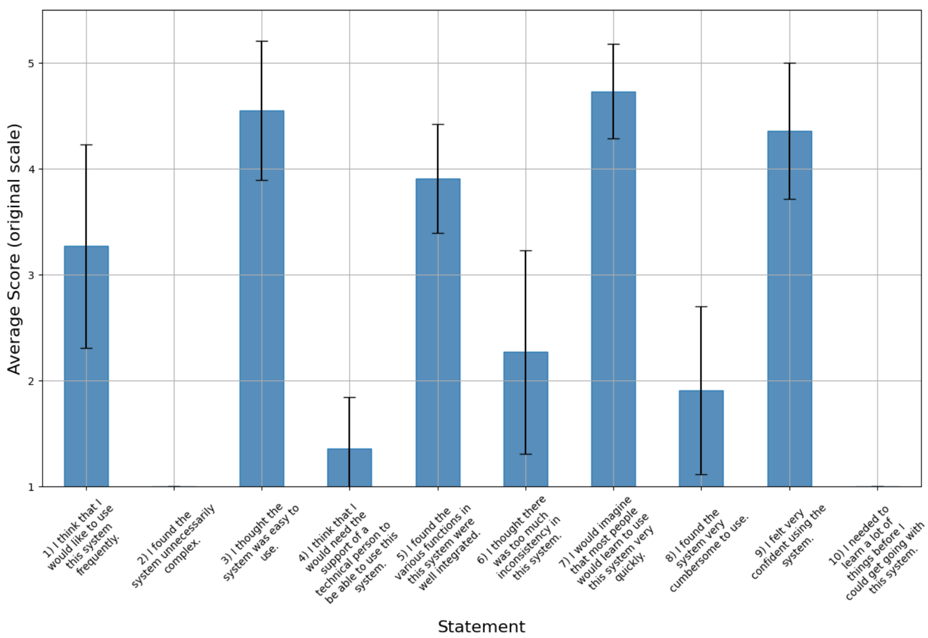


Figure 6.2: Average score (original scale) for each of the 10 SUS statements with belonging standard deviation.

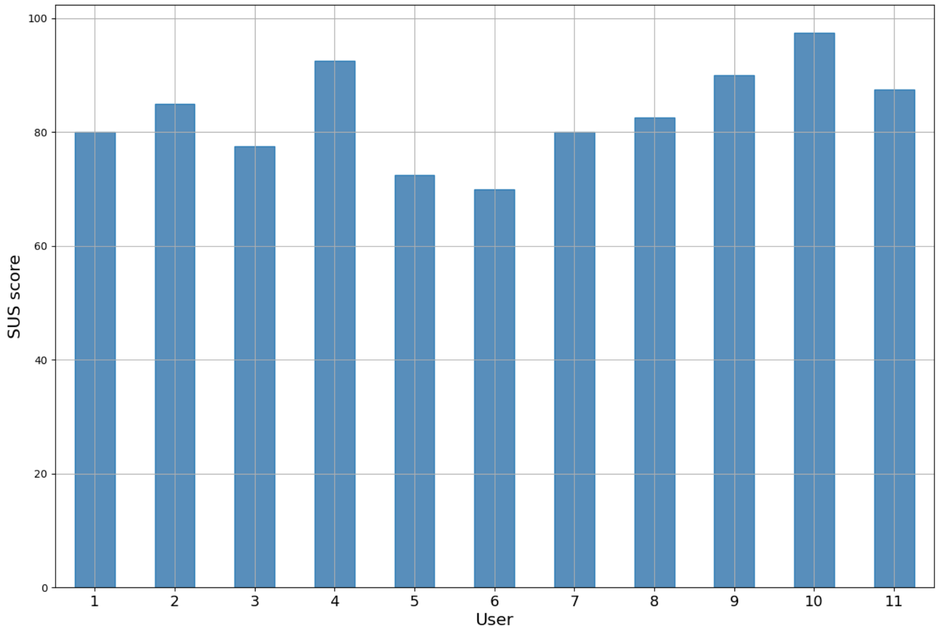


Figure 6.3: The individual SUS score for each of the 11 test participants.

6.4.2 Qualitative results

In addition to the ten statements, the SUS form has an additional section where users can provide additional comments regarding the system. Nine of the eleven test participants chose to utilize this section to add some supplementary comments regarding how they experienced the system. A visualized summary of the addressed themes is found in figure 6.4.

A comment that was repeated was that several users experienced the spawning of the fruit to be more biased to one side of the skeleton than the other. This resulted in one arm being activated more than the other. As the fruits were set to spawn randomly with equal probability of spawning on either side of the user, this was somewhat surprising feedback. Some users also reported they experienced the fruit to sometimes be out of reach. This also ties in with the spawn settings. It was very useful that the SUS form was able to pick up this inconsistency, so the system can be adjusted accordingly. Positive comments were given to the game scenery and level of immersion the game was able to provide. Participants found the nature themed scenery pleasant. Three participants also commented they found the game to be more fun when played at the more difficult levels. Although only three people made this comment explicitly in the SUS form, this was a response several more participants gave verbally after trying the game at different difficulty levels. Regarding the audio, two people made concrete remarks that they found the audio pleasant. One person suggested that there could be variation in the audio, so that different fruits had different spawn sounds. Another participant remarked that an idea would be for the audio to be used as a more precise indicator for where the fruit would spawn. However, given the small spawn range of the spawned fruits, this would likely not be very effective. Five participants reported experiencing some inconsistency with the pose tracking glitching and being randomly unresponsive. Having the pose estimation model not working properly with the system would be an unsustainable and unacceptable alternative if the system was to engage with stroke survivors in post-stroke rehabilitation.

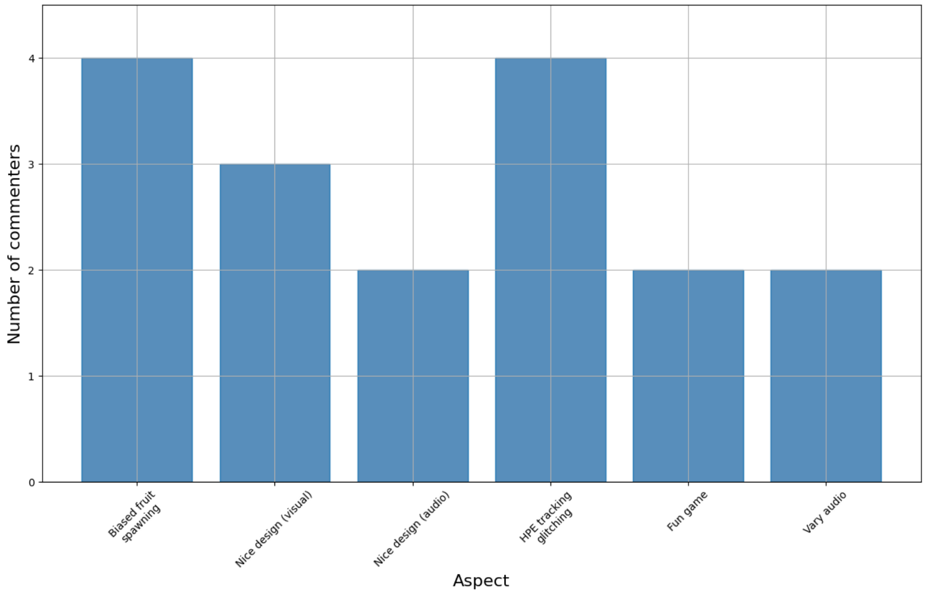


Figure 6.4: Aspects of the system addressed by participants in the written feedback

Chapter 7

Discussions

7.1 Design approach

The project followed a methodology for designing therapeutic XR exergames. The methodology provided a guideline for how to begin the process of designing and developing for serious rehabilitation games. The steps of the framework were clearly defined, making it intuitive to understand how to proceed with the instructions of the different stages. Gaining access to specific resources to follow the framework like the VR game canvas and the miro-boards for concept ideation provided an additional element of practicality.

7.1.1 Choice of tools

Unity provided a great deal of useful documentation and tutorials through its manual [66] and chat forum [65]. I was able to ask questions in the chat forum and be greeted by a community of developers. The accessibility of information was very practical and made the build phase easier to approach. The engine provided a selection of free assets ranging from 3D models to audio clips, and integrated assets and packages into the project seamlessly.

The HTC Vive worked well for creating immersive VR experience for the player. A limitation of the HTC Vive headset is that it requires several wires to connect to a computer. It requires connection to a display port, a USB port, and a power outlet simultaneously. Using a HMD with a wireless connection could simplify the setup process. For example, one challenged encountered in this project, was that the development laptop in use did not possess a display port outlet, and it was tricky to find an available converter. The absence of wires would also remove the safety concern of the user stumbling on the wires.

7.2 Human pose estimation with BlazePose

Several participants experienced some glitching with the movements of the avatar, causing a perceived lack of complete control. Feeling as they were not always in complete control of the avatar negatively affected the immersiveness of the experience. During the development phase of the game, tests were occasionally run using an recorded video input instead of webcam input, as it meant it wasn't necessary to fix the technical setup of the exergame each time. When running the game with a person on a recorded video controlling the avatar, there were no noticeable glitching issues and the control of the avatar appeared seamless. Faulty connection between the avatar and the user in the real-life setting can be due to several reasons. If the user temporarily leaves the camera frame, the model is unable to track all key points necessary for correct pose estimation. This issue was attempted mitigated by ensuring that the participants were fully visible in the frame of the webcam before testing began. Similarly, issues can be caused if the color of the participants attire was too close to the background color, making it difficult for the model to identify all key tracking points. The fact that the recorded video input was able to control the avatar so seamlessly makes it plausible that the issue could be related to the physical test conditions rather than the mapping between the BlazePose model and the user's avatar. If the issue should reside in the mapping it is likely related to the combination of the HMD and the BlazePose model, as this, besides the physical environment, is the aspect that differs between using a recorded video as input and webcam input. Having identified these potential origins to the issue can be of benefit for future iterations of the systems. The implementation of the humanoid avatar for VR without controllers was challenging due to a lack of documentation regarding using HPE models to control humanoid avatars in VR. The examples found regarding this were either without the use of VR [40], or with VR but with the use of controllers for mapping the hands [46].

7.3 Limitations

Throughout the completion of this project several limitations regarding various aspects were identified.

Section 3: Related Work clarified that there, as of yet, are few examples of BlazePose being utilized in rehabilitation settings. Thus, the main novel aspect of this project, was the use of HPE in a VR post-stroke rehabilitation game. Therefore, few previous works and experiences were available to reference and learn from, causing an additional challenge. Not all stroke victims would be eligible for a VR stroke rehabilitation system [57]. Most of the research that has been conducted on VR schemes for stroke recovery mainly focuses on users who have suffered strokes that fall under the classification "mild to moderate". If these systems would be applicable for victims of more severe strokes is still uncertain, but can be explored through future research [57]. This could potentially limit the scope of post-stroke rehabilitation systems.

This project narrowed its research to mainly concern upper-limb motor rehabilitation exercises fairly early in its lifespan. The exclusion of lower-limb and mobility exercises limits the range of types of motor recovery the system is applicable for. Overcoming this limitation would open up the development of exergames for motor function recovery in stroke survivors to encompass all acquired motor disabilities. To extend the system to exercises beyond the upper-limb scope, the safety concern of lower-limb and balance exercises would need to be addressed.

7.4 Evaluation

The participants of the usability test procedures did not reflect the real intended user group for this system. The decision to test the system on healthy, easily available users was made based on a range of considerations. This project was a proof of concept exploring the combination of HPE and VR in exergames for post-stroke motor rehabilitation, and was constrained to a short time frame. Getting the prototype to a stage where the exercise impact of the application could be properly evaluated was not deemed realistic, which would somewhat defeat the point of including real test users. Also, getting NSD applications approved for using patients in user tests is more complicated than getting approval for healthy participants. All participants of the usability tests reside in a relatively young age range, future testing of the system should be expanded to include users of older age. This would be significant as there has been proven to be a slight relationship between the age of participants and results of the SUS [2]. Older participants score somewhat lower on the SUS than younger participants. As the probability of having a stroke increases after surpassing the age of 45, conducting tests with an older age group would be a necessary step in eventual future iterations of system testing.

Participants for the usability test were recruited from my network of acquaintances. Like myself, many of them have a technical background. This leaves statement four of the SUS "I think that I would need the support of a technical person to be able to use this system" [B] somewhat obsolete, as several of the participants by some measure can be considered "technical people" and are assumed to possess a higher degree of technical competency and understanding than the average person. Although the system achieved a good score from the SUS assessment, on must keep in mind that the SUS is better at identifying system flaws than it is at assessing when a system is good enough. With this and the qualitative feedback of the participant's in mind, one must assume there is still much to that can be improved in this system.

7.5 Further work

This thesis project can be described as a proof of concept to explore the realization of an exergame for users in post-stroke motor function recovery that combines virtual reality and pose estimation technologies. The final aim of further exploring the topic of motor rehabilitation exergames for stroke survivors would be to create technically sound prototypes with validated therapeutic outcomes that can be deployed to end-users.

Another natural progression for this project would be the involvement of patients in post-stroke recovery and physical therapists. The inclusion of relevant stakeholders can offer valuable insights into how the game can be expanded or rebuilt to even better meet the needs of this user group. This can be achieved by involving stakeholders in the definition and validation stages of the game design framework. Following a rigorous inclusion criteria to identify test participants that fit in the system's intended user group could provide invaluable insights into how to better design games for stroke survivors by getting feedback directly from the people who are affected.

The current game solely focuses on upper-limb movements mainly pertaining to training of the arms and shoulders. However, there exists a myriad of motor function rehabilitation exercises that could be relevant to incorporate into exercise games for stroke survivors. Expanding the amount of implemented exercises would be a natural step forward. The SUS was an appropriate tool for measuring the quality of the project in terms of usability while the project was in a preliminary stage. If the projects was to progress into a more serious development phase, additional test procedures should be explored as extensions of the current testing scheme. Using the SUS can still be relevant to get an indication of future iterations of the project in terms of usability, but should not be the only quality measure the system must undergo. For example, expanding the test regime to include schemes for also measuring the therapeutic outcomes of the game could greatly strengthen the validity of the game in terms of therapeutic efficacy. One suggestion is using the Fugl-Meyer Assessment (FMA) [22] tool for measuring motor impairment and improvement following a stroke to integrate evaluation of the rehabilitation outcome into a test regime. Currently, the system provides the user with feedback regarding if they have completed an exercise or not. An interesting avenue to explore would be to also integrate feedback on *how* the users execute the rehabilitation exercises. Being able to provide feedback regarding the technical execution of the exercises and how the user could potentially improve, could be beneficial to further advance the system in terms of therapeutic validity. This would be an exciting avenue to explore in further iterations of this system.

There was an intention to incorporating rhythmic elements into the exergame as this was identified as a novel aspect of rehabilitative exergames that is still in an exploratory phase. While auditory elements were incorporated to coincide with the spawning of fruit elements, the auditory elements can be further expanded to make the rhythmic aspects a more wholesome and well-integrated part of the exergame. This could be achieved by further researching the most beneficial way to combine music with motor function rehabilitation.

Several test participants identified irregularities regarding the motion capture technique the system utilizes, BlazePose. Further research is necessary to understand how these limitations can be overcome so users can have a more smooth experience. The project showed that the model can be integrated into an exergame with acceptable quality in terms of usability, so abandoning the technology altogether should not be necessary. However, the feedback indicates that further action needs to be taken for the system to be in an acceptable state in terms of motion capture. Implementing a scaling procedure could be an interesting path to explore as it could potentially match the avatar better to each individual user. The advancement of machine learning and pose estimation technologies could also contribute to improved performance. Testing out the use of other motion capture technologies could be an option if the limitations prove difficult to overcome. This prototype and report is a preliminary result to explore the concept of combining HPE and VR in rehabilitative exergames. Future iterations of the system should put more focus on implementing features that lead to user's experiencing increased motivation and adherence in their recovery. Focusing on enriching the high-level requirements that were identified in the preliminary study, namely meaningful feedback, user safety, user-centric design, rhythmic elements and task specificity, would make the system capable of creating an even more motivating user experience.

Chapter 8

Conclusions

In this project a prototype of a virtual reality exergame utilizing the pose estimation algorithm BlazePose for post-stroke upper-limb motor rehabilitation was developed.

This report documents the process to create the prototype. It also provides an account of the motivation behind conducting research on this topic, relevant background theory to understand the context of the project and a related work section detailing projects that have been conducted on similar themes. Additionally, an account of the methodology utilized for the design- and development phase of the prototype, the resulting prototype, a description of the chosen evaluation scheme and results from the user testing are given. This is followed by a discussion regarding the various aspects of the project, as well as a conclusion, answering the research questions which were introduced.

Following the serious game design framework of Støen and Fridheim [59] provided a methodological approach to developing serious VR rehabilitation games. The methodology was clear to follow, could be executed with few external resources and provided results within a very reasonable time-frame. This methodology was successfull in guiding the game design process of a serious rehabilitation games. The process results in a solidly documented game design process that can be observed and analyzed by interested parties and potential stakeholders.

Challenges and opportunities relating to the use of pose estimation with BlazePose for a virtual reality post-stroke motor rehabilitation exergame were identified. Due to the novelty of pose estimation technologies there is a scarcity of similarly themed projects from which one can take learning and inspiration. Several of the evaluation participants reported that their experience with the tracking technology was suboptimal. Even so, the final prototype of the system was deemed acceptable in terms of usability following the criteria of the SUS. Although issues were identified there is much potential in the use of HPE in rehabilitation interventions, and further research is required to determine what steps can be taken to make HPE motion capture more optimal for exergames. Opportunities that were identified in using HPE for this project was the accessibility of the technology. BlazePose was relatively uncomplicated to integrate into Unity and the webcam was the only hardware requirement besides the laptop necessary for the pose estimation aspect of the system.

The key technical considerations required for real-time human pose estimation integration into a VR exergame for stroke rehabilitation are the implementation of an avatar that maps the keypoints of the pose estimation, the implementation of gameplay that rely on the pose estimation/ avatar movements, and adjustments required to integrate an HPE model that relies on keypoints in the facial area with the use of a HMD.

References

[1] Aritz Badiola-Bengoa and Amaia Mendez-Zorrilla. 'A systematic review of the application of camera-based human pose estimation in the field of sport and physical exercise'. In: *Sensors* 21.18 (2021), p. 5996.

[2] Aaron Bangor, Philip Kortum and James Miller. 'Determining what individual SUS scores mean: Adding an adjective rating scale'. In: *Journal of usability studies* 4.3 (2009), pp. 114–123.

[3] Aaron Bangor, Philip T. Kortum and James T. Miller. 'An Empirical Evaluation of the System Usability Scale'. In: *International Journal of Human–Computer Interaction* 24.6 (29th July 2008). Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/10447310802205776, pp. 574–594. ISSN: 1044-7318. DOI: 10.1080/10447310802205776. URL: https://doi.org/10.1080/10447310802205776 (visited on 8th May 2023).

[4] Valentin Bazarevsky and Ivan Grishchenko. *On-Device, Real-Time Body Pose Tracking with MediaPipe BlazePose*. https://ai.googleblog.com/2020/08/on-device-real-time-body-pose-tracking.html. 2020. (Visited on 12th May 2023).

[5] Valentin Bazarevsky et al. 'Blazepose: On-device real-time body pose tracking'. In: *arXiv preprint arXiv:2006.10204* (2020).

[6] Francesco Bellotti et al. 'Assessment in and of serious games: an overview'. In: *Advances in human-computer interaction* 2013 (2013), pp. 1–1.

[7] Miguel Borges et al. 'HTC vive: Analysis and accuracy improvement'. In: *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2018, pp. 2610–2615.

[8] Steven Chen and Richard R. Yang. *Pose Trainer: Correcting Exercise Posture using Pose Estimation*. 21st June 2020. DOI: 10.48550/arXiv.2006.11718. arXiv: 2006.11718[cs]. URL: http://arxiv.org/abs/2006.11718 (visited on 10th June 2023).

[9] HTC Corporation. *VIVE Pro HMD Setup guide*. 2023. URL: https://www.vive.com/us/setup/vive-pro-hmd/ (visited on 10th June 2023).

[10] Microsoft Corporation. *Azure Kinect DK: Build for mixed reality using AI sensors*. 2023. URL: https://azure.microsoft.com/en-gb/products/kinect-dk/ (visited on 9th June 2023).

[11] Microsoft Corporation. *LifeCam Studio*. 2023. URL: https://www.microsoft.com/en/accessories/products/webcams/lifecam-studio?activetab=pivot:overviewtab (visited on 7th June 2023).

[12] Valve Corporation. *SteamVR*. 2023. URL: https://www.steamvr.com/en/ (visited on 3rd June 2023).

[13] creativeIKEP. *BlazePoseBarracuda*. Version 1.3.0. Sept. 2022. URL: https://github.com/creativeIKEP/BlazePoseBarracuda.

[14] JH Crosbie et al. 'Virtual reality in stroke rehabilitation: still more virtual than real'. In: *Disability and rehabilitation* 29.14 (2007), pp. 1139–1146.

[15] Paul Dempsey. 'The teardown: HTC Vive VR headset'. In: *Engineering & Technology* 11.7-8 (2016), pp. 80–81.

[16] Nancy E. Mayo et al. 'Disablement following stroke'. In: *Disability and rehabilitation* 21.5-6 (1999), pp. 258–268.

[17] Hao-Shu Fang et al. 'Alphapose: Whole-body regional multi-person pose estimation and tracking in real-time'. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2022).

[18] Bruno Ferreira and Paulo Menezes. *An Adaptive Virtual Reality-Based Serious Game for Therapeutic Rehabilitation*. Pages: 63-71. International Association of Online Engineering, 8th Apr. 2020. URL: https://www.learntechlib.org/p/217960/ (visited on 6th June 2023).

[19] Free Software Foundation. *GNU General Public License*. 2023. URL: http://www.gnu.org/licenses/gpl.html (visited on 1st June 2023).

[20] The Apache Software Foundation. *Apache License*. 2004. URL: https://www.apache.org/licenses/LICENSE-2.0 (visited on 4th June 2023).

[21] Francesca Gimigliano and Stefano Negrini. 'The world health organization" rehabilitation 2030: a call for action"'. In: *European Journal of Physical and Rehabilitation Medicine* 53.2 (2017), pp. 155–168.

[22] David J Gladstone, Cynthia J Danells and Sandra E Black. 'The Fugl-Meyer assessment of motor recovery after stroke: a critical review of its measurement properties'. In: *Neurorehabilitation and neural repair* 16.3 (2002), pp. 232–240.

[23] Ivan Grishchenko et al. *BlazePose GHUM Holistic: Real-time 3D Human Landmarks and Pose Estimation*. 23rd June 2022. DOI: 10.48550/arXiv.2206.11678. arXiv: 2206.11678[cs]. URL: http://arxiv.org/abs/2206.11678 (visited on 7th Mar. 2023).

[24] Graeme J Hankey. 'Stroke'. In: *Lancet* 389.10069 (2017), pp. 641–654.

[25] Marianne Hernholm. 'Exploring VR exergames for post-stroke motor function rehabilitation'. Department of Computer Science, Norwegian University of Science and Technology, 2022.

[26] Marianne Hernholm. *RehabExergame*. June 2023. URL: https://github.com/marholm/RehabExergame.

[27] Ginés Hidalgo et al. *OpenPose: The first real-time multi-person system to jointly detect human body, hand, facial, and foot keypoints*. 2023. URL: https://cmu-perceptual-computing-lab.github.io/openpose/web/html/doc/.

[28] I-Ping Hsueh et al. 'A Simplified Stroke Rehabilitation Assessment of Movement Instrument'. In: *Physical Therapy* 86.7 (1st July 2006), pp. 936–943. ISSN: 0031-9023. DOI: 10.1093/ptj/86.7.936. URL: https://doi.org/10.1093/ptj/86.7.936 (visited on 6th Oct. 2022).

[29] The Khronos Group Inc. *OpenXR*. 2023. URL: https://www.khronos.org/OpenXR/ (visited on 4th May 2023).

[30] Walter Johnson et al. 'Stroke: a global response is needed'. In: *Bulletin of the World Health Organization* 94.9 (2016), p. 634.

[31] Yeong Wook Kim. 'Update on Stroke Rehabilitation in Motor Impairment'. In: *Brain & Neurorehabilitation* 15.2 (2022).

[32] Kinetisense. *What Is a Functional Movement Screening?* 2023. URL: https://www.kinetisense.com/a-functional-movement-screen/ (visited on 9th June 2023).

[33] Jyothsna Kondragunta and Gangolf Hirtz. 'Gait parameter estimation of elderly people using 3D human pose estimation in early detection of dementia'. In: *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE. 2020, pp. 5798–5801.

[34] Audrius Kulikajevas et al. 'Exercise abnormality detection using BlazePose skeleton reconstruction'. In: *Computational Science and Its Applications–ICCSA 2021: 21st International Conference, Cagliari, Italy, September 13–16, 2021, Proceedings, Part V 21*. Springer. 2021, pp. 90–104.

[35] Fedwa Laamarti, Mohamad Eid and Abdulmotaleb El Saddik. 'An overview of serious games'. In: *International Journal of Computer Games Technology* 2014 (2014), pp. 11–11.

[36] Kate E Laver et al. 'Virtual reality for stroke rehabilitation'. In: *Cochrane database of systematic reviews* 11 (2017).

[37] Gwyn N Lewis and Juliet A Rosie. 'Virtual reality games for movement rehabilitation in neurological conditions: how do we meet the needs and expectations of the users?' In: *Disability and rehabilitation* 34.22 (2012), pp. 1880–1886.

[38] James R. Lewis. 'The System Usability Scale: Past, Present, and Future'. In: *International Journal of Human–Computer Interaction* 34.7 (3rd July 2018). Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/10447318.2018.1455307, pp. 577–590. ISSN: 1044-7318. DOI: 10.1080/10447318.2018.1455307. URL: https://doi.org/10.1080/10447318.2018.1455307 (visited on 8th May 2023).

[39] C Lindmeier and A Brunier. 'WHO: number of people over 60 years set to double by 2050; major societal changes required'. In: *World Health Organization* (2015).

[40] Digital Standard Co. LTD. *ThreeDPoseUnityBarracuda*. 2021. URL: https://github.com/digital-standard/ThreeDPoseUnityBarracuda (visited on 3rd Mar. 2023).

[41] Camillo Lugaresi et al. *MediaPipe: A Framework for Building Perception Pipelines*. 14th June 2019. arXiv: 1906.08172[cs]. URL: http://arxiv.org/abs/1906.08172 (visited on 13th May 2023).

[42] Shanthi Mendis. 'Stroke disability and rehabilitation of stroke: World Health Organization perspective'. In: *Int J stroke* 8.1 (2013), pp. 3–4.

[43] Farouk Messaoudi, Gwendal Simon and Adlen Ksentini. 'Dissecting games engines: The case of Unity3D'. In: *2015 international workshop on network and systems support for games (NetGames)*. IEEE. 2015, pp. 1–6.

[44] Evan Paul Miller. 'An Experimental Evaluation of Methods for 3D Pose Estimation with Avatar Animation Using A Single RGB Camera'. MA thesis. California State Polytechnic University, 2021.

[45] Timothy H Murphy and Dale Corbett. 'Plasticity during stroke recovery: from synapse to behaviour'. In: *Nature reviews neuroscience* 10.12 (2009), pp. 861–872.

[46] Maciek Odrowaz. *FullBodyPoseEstimation*. Dec. 2022. URL: https://github.com/maceq687/FullBodyPoseEstimation.

[47] Maciej Odrowaz-Sypniewski and Sofia Lamda. 'Implementing full-body pose estimation for Virtual Reality applications: A brief comparison among attainable tools'. MA thesis. Aalborg University, 2022.

[48] Yoonsin Oh and Stephen Yang. 'Defining exergames & exergaming'. In: *Proceedings of meaningful play* 2010 (2010), pp. 21–23.

[49] World Health Organization. *Ageing and health*. 2022. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health (visited on 6th June 2023).

[50] World Health Organization. *Rehabilitation*. 2023. URL: https://www.who.int/news-room/fact-sheets/detail/rehabilitation (visited on 12th May 2023).

[51] World Health Organization. *World report on ageing and health*. World Health Organization, 2015.

[52] Ainun Syarafana Binti Pauzi et al. 'Movement estimation using mediapipe blazepose'. In: *Advances in Visual Informatics: 7th International Visual Informatics Conference, IVIC 2021, Kajang, Malaysia, November 23–25, 2021, Proceedings 7*. Springer. 2021, pp. 562–571.

[53] Michele Pirovano et al. 'Exergaming and rehabilitation: A methodology for the design of effective and safe therapeutic exergames'. In: *Entertainment Computing* 14 (2016), pp. 55–65.

[54] Pure Poly. *Free Low Poly Nature Forest*. 2022. URL: https://assetstore.unity.com/packages/3d/environments/landscapes/free-low-poly-nature-forest-205742 (visited on 1st Apr. 2023).

[55] Google Research. *Pose landmark detection guide*. https://developers.google.com/mediapipe/solutions/vision/pose_landmarker/. May 2023. (Visited on 20th Mar. 2023).

[56] Steven R. Rick et al. 'NeuroPose: geriatric rehabilitation in the home using a webcam and pose estimation'. In: *Proceedings of the 24th International Conference on Intelligent User Interfaces: Companion*. IUI '19. New York, NY, USA: Association for Computing Machinery, 16th Mar. 2019, pp. 105–106. ISBN: 978-1-4503-6673-1. DOI: 10.1145/3308557.3308682. URL: https://dl.acm.org/doi/10.1145/3308557.3308682 (visited on 10th June 2023).

[57] Gustavo Saposnik, Mindy Levin and Stroke Outcome Research Canada (SORCan) Working Group. 'Virtual reality in stroke rehabilitation: a meta-analysis and implications for clinicians'. In: *Stroke* 42.5 (2011), pp. 1380–1386.

[58] Nikolaos Sarafianos et al. '3d human pose estimation: A review of the literature and analysis of covariates'. In: *Computer Vision and Image Understanding* 152 (2016), pp. 1–20.

[59] Anders B. Støen and Magnus L. Fridheim. 'VRehab - Designing for VR-based serious games in neurorehabilitation'. Master thesis. Norwegian University of Science and Technology, June 2021.

[60] Rem storms. *Low Poly Fruit Pickups*. 2018. URL: https://assetstore.unity.com/packages/3d/props/food/low-poly-fruit-pickups-98135 (visited on 1st Apr. 2023).

[61] Polytope Studio. *Lowpoly Medieval Skeleton*. 2020. URL: https://assetstore.unity.com/packages/3d/characters/humanoids/fantasy/lowpoly-medieval-skeleton-free-medieval-fantasy-series-181883 (visited on 1st Apr. 2023).

[62] Sandeep K. Subramanian et al. 'Arm Motor Recovery Using a Virtual Reality Intervention in Chronic Stroke: Randomized Control Trial'. In: *Neurorehabilitation and Neural Repair* 27.1 (Jan. 2013), pp. 13–23. ISSN: 1545-9683, 1552-6844. DOI: 10.1177/1545968312449695. URL: http://journals.sagepub.com/doi/10.1177/1545968312449695 (visited on 6th June 2023).

[63] Unity Technologies. *Manual: Introduction to Barracuda*. Oct. 2020. URL: https://docs.unity3d.com/Packages/com.unity.barracuda@1.0/manual/index.html.

[64] Unity Technologies. *Unity Asset Store*. 2023. URL: https://assetstore.unity.com/ (visited on 28th May 2023).

[65] Unity Technologies. *Unity Forum*. 2023. URL: https://forum.unity.com/ (visited on 9th June 2023).

[66] Unity Technologies. *Unity User Manual 2021.3 (LTS)*. 2023. URL: https://docs.unity3d.com/Manual/UnityManual.html (visited on 28th May 2023).

[67] Mateus Trombetta et al. 'Motion Rehab AVE 3D: A VR-based exergame for post-stroke rehabilitation'. In: *Computer methods and programs in biomedicine* 151 (2017), pp. 15–20.

[68] Elise Klæbo Vonstad et al. 'Comparison of a deep learning-based pose estimation system to marker-based and kinect systems in exergaming for balance training'. In: *Sensors* 20.23 (2020), p. 6940.

[69] Guangjun Wang et al. 'Design and evaluation of an exergame system to assist knee disorders patients' rehabilitation based on gesture interaction'. In: *Health Information Science and Systems* 10.1 (26th Aug. 2022), p. 20. ISSN: 2047-2501. DOI: 10.1007/s13755-022-00189-5. URL: https://doi.org/10.1007/s13755-022-00189-5 (visited on 6th June 2023).

[70] Séamas Weech, Sophie Kenny and Michael Barnett-Cowan. 'Presence and cybersickness in virtual reality are negatively related: a review'. In: *Frontiers in psychology* 10 (2019), p. 158.

[71] Patrice L Weiss et al. 'Virtual reality in neurorehabilitation'. In: *Textbook of neural repair and rehabilitation* 51.8 (2006), pp. 182–97.

[72] Charles DA Wolfe. 'The impact of stroke'. In: *British medical bulletin* 56.2 (2000), pp. 275–286.

[73] Erwin Wu and Hideki Koike. 'Futurepose-mixed reality martial arts training using real-time 3d human pose forecasting with a rgb camera'. In: *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE. 2019, pp. 1384–1392.

[74] Jingming Xie. 'Research on key technologies base Unity3D game engine'. In: *2012 7th international conference on computer science & education (ICCSE)*. IEEE. 2012, pp. 695–699.

[75] Jeffrey Yim and T. C. Nicholas Graham. 'Using games to increase exercise motivation'. In: *Proceedings of the 2007 conference on Future Play - Future Play '07*. Toronto, Canada: ACM Press, 2007, p. 166. ISBN: 978-1-59593-943-2. DOI: 10.1145/1328202.1328232. URL: http://portal.acm.org/citation.cfm?doid=1328202.1328232 (visited on 5th Dec. 2022).

[76] Ce Zheng et al. 'Deep learning-based human pose estimation: A survey'. In: *arXiv preprint arXiv:2012.13392* (2020).

Appendix A

Information and consent form for test participants

A.1 Information form

The consent form asks the user to explicitly consent to participation in the user-test and anonymously answering the system usability scale (SUS)-schema after completing the user test.The consent form asks that the information obtained from the SUS-schema can be used in my master thesis.

Exergames for rehabilitation
This is an inquiry about participation in a master project by a 5th year M.Sc. Computer Science student from the Department of Computer Science at NTNU. In this letter you will be given information about the purpose of this project and what your participation will involve.

Purpose of the project
The purpose of the project is to research the use of exercise games with extended reality (XR) and human pose estimation (HPE) aspects for patients in sub-acute phases of post-stroke recovery. The goal is to identify if systems utilizing these features are technically viable, and if they can have a positive impact in the rehabilitation process of this user group. The system will focus on motor rehabilitation.

What does participation involve for you?
Choosing to participate in this project is voluntary and will involve testing the created system and its functionalities. You will play an exergame consisting of completing two simple exercises while wearing a head mounted display (HMD)-headset. Your movements will be tracked through a web-cam that feeds the input to a pose estimation algorithm called BlazePose. After testing the system you will be asked to fill out the standardized system usability scale (SUS)- schema. This form consists of 10 questions that you will give a score of 1 (strongly disagree) to 5 (strongly agree). The information from the form is anonymous and will not include any details that can identify you individually.

Participation is voluntary
Participation in the project is voluntary. There will be no negative consequences for you if you choose not to participate.

Your personal privacy - how we will store and use your personal data
The publications will not contain any uniquely identified data, although your status as a student/ profession, gender and age will be published. Also it will be mentioned that you do not have any severe motor disabilities. This is due to the fact that we will specify not using actual users with motor disabilities, but that our participants did not have motor-function challenges.

What will happen to your personal data at the end of the research project?
The project is scheduled to end in June 2023. At that time, all records of your participation will be

deleted from all digital systems. The data published in the paper will still be accessible, although anonymised as mentioned previously.

Your rights
So long as you can be identified in the collected data, you have the right to: access the personal data that is being processed about you request that incorrect personal data about you is corrected/rectified receive a copy of your personal data (data portability), and send a complaint to the Data Protection Officer or The Norwegian Data Protection Authority regarding the processing of your personal data The collected data will not contain information that you can be identified by.

What gives us the right to process your personal data?
We will process your personal data based on your consent. Based on an agreement with NTNU's Department of Computer Science, The Norwegian Centre for Research Data (NSD) has assessed that the processing of personal data in this project is in accordance with data protection legislation.

Where can I find out more?
If you have questions about the project, or want to exercise your rights, contact:

- Marianne Hernholm: ********@stud.ntnu.no
- Norwegian Centre for Research Data (NSD)

A.2 Consent form

Statement of consent

1. Participate in a user-test of the exergame wearing a HMD and being tracked by the BlazePose algorithm.
2. Answer the SUS-schema anonymously regarding my experience testing the exergame.
3. The researcer processing the data from the SUS-schema and including the results in their master thesis.
4. My personal data to be processed until the end date of the project, approx. June 2023.

Do you consent to participation in the project?
Yes
No

Date of consent

Do you have any serious motor function disabilities?
Yes
No

Name

Date of Birth

Gender
Female
Male
Other(s)

E-mail

Appendix B

System Usability Scale

1. I think that I would like to use this system frequently.
2. I found the system unnecessarily complex.
3. I thought the system was easy to use.
4. I think that I would need the support of a technical person to be able to use this system.
5. I found the various functions in this system were well integrated.
6. I thought there was too much inconsistency in this system.
7. I would imagine that most people would learn to use this system very quickly.
8. I found the system very cumbersome to use.
9. I felt very confident using the system.
10. I needed to learn a lot of things before I could get going with this system.

Additional thoughts/ comments:

Appendix C

The VR Game Canvas

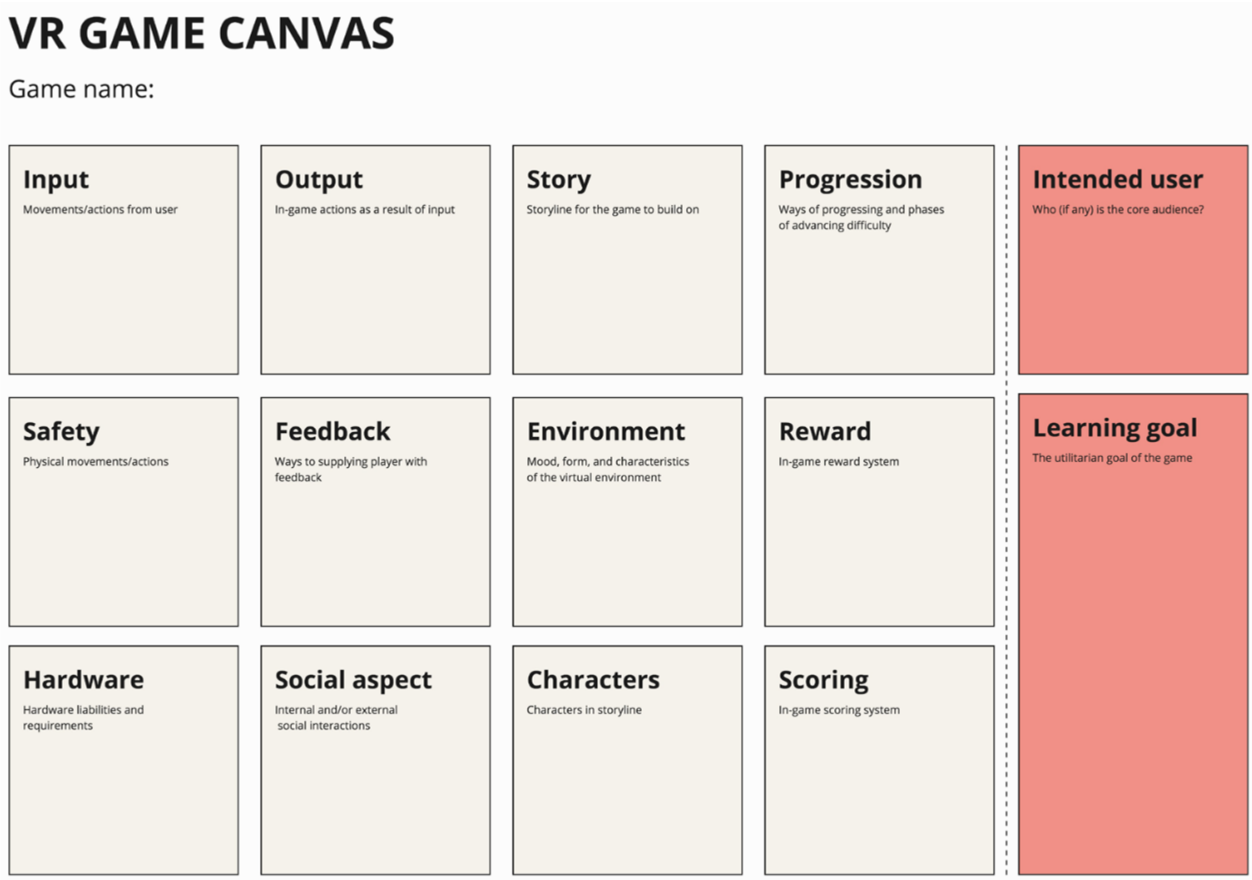


Figure C.1: The VR Game Canvas [59]

Appendix D

Generated game concepts

D.1 Maestro concept

Maestro concept shown in D.1.

D.2 Volleyball concept

Volleyball concept shown in D.2.

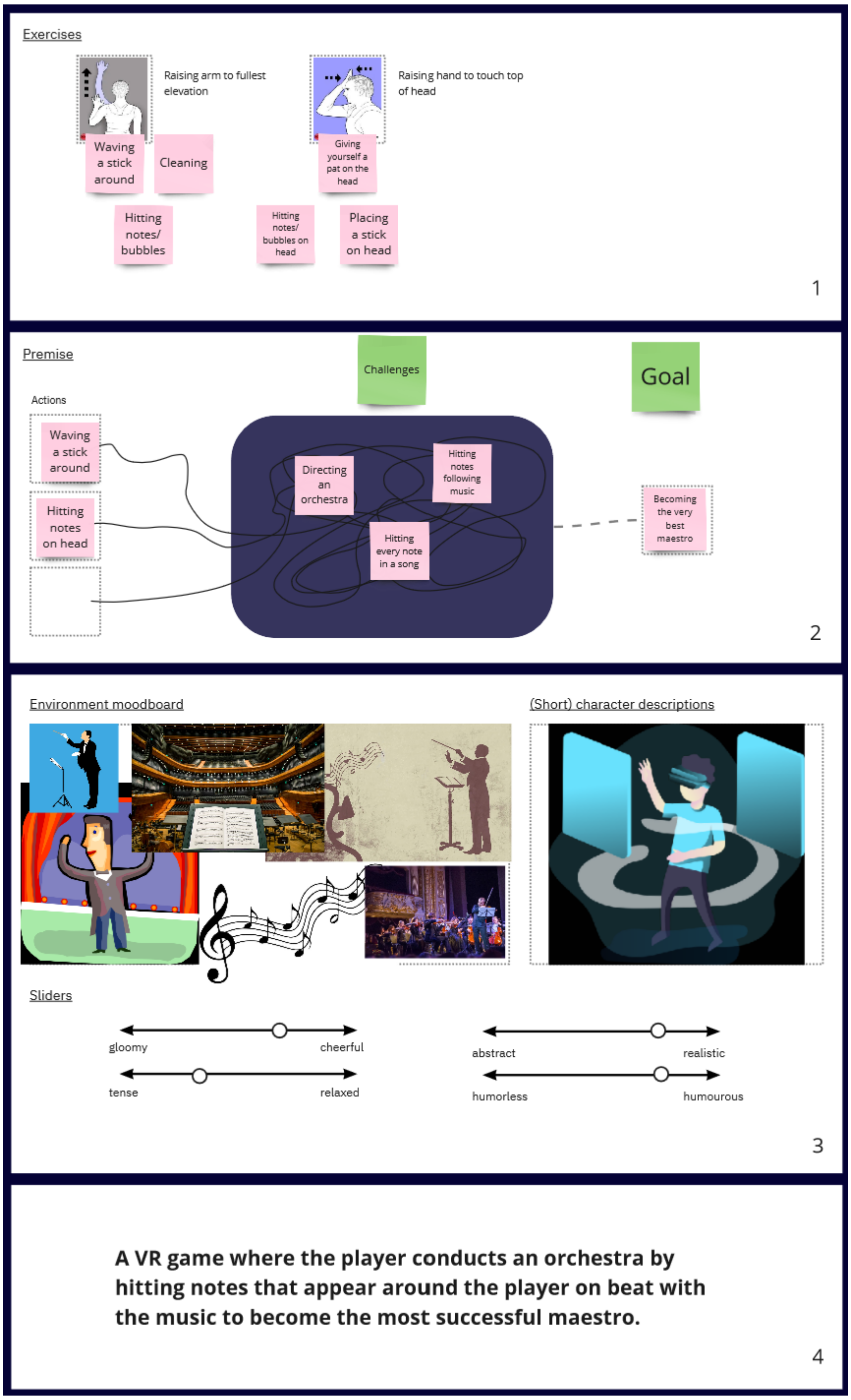


Figure D.1: Maestro game concept generated through the Støen and Fridheim [59] framework. All images accessed through creative commons license.

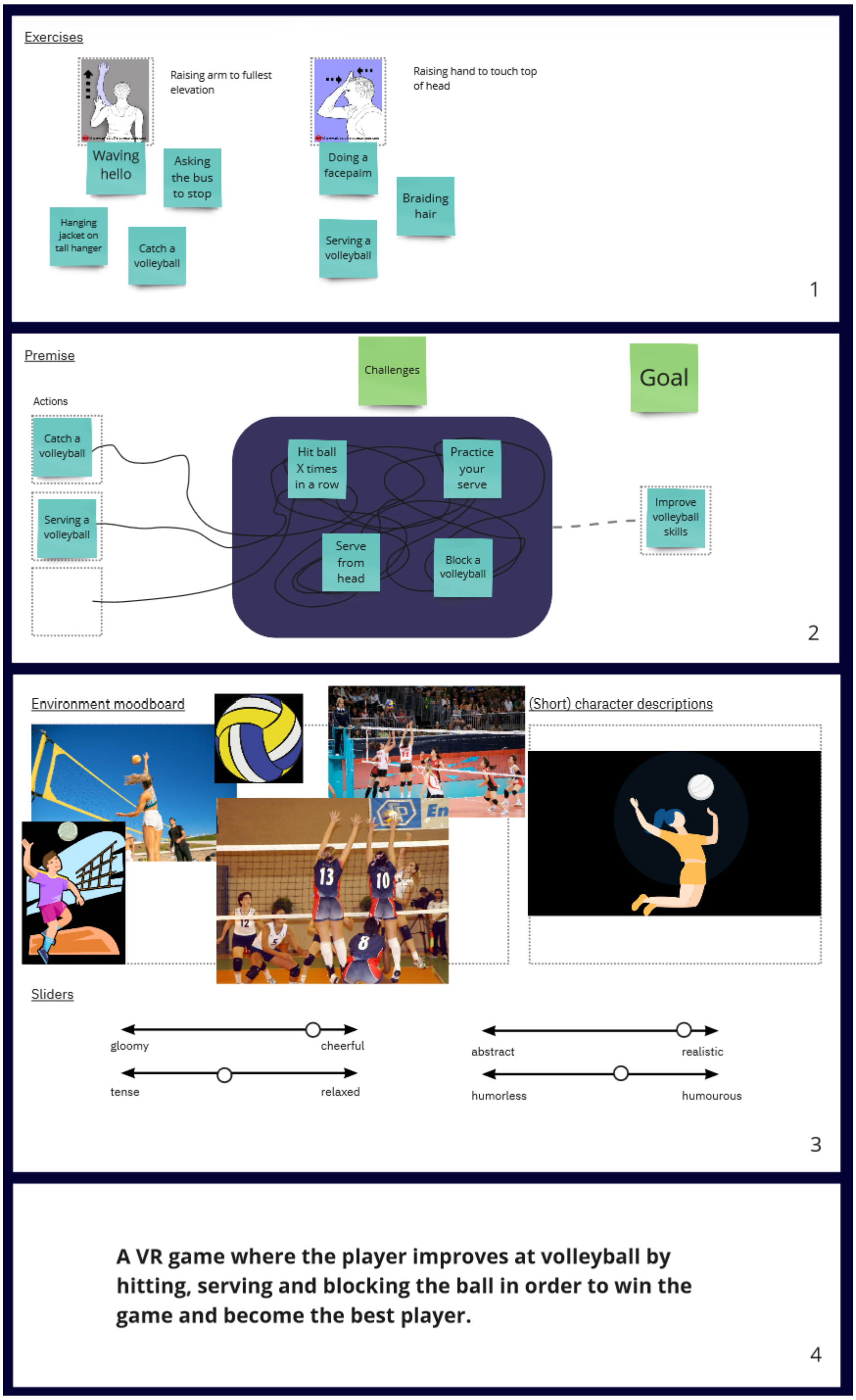


Figure D.2: Volleyball game concept generated through the Støen and Fridheim [59] framework. All images accessed through creative commons license.

NTNU
Kunnskap for en bedre verden