Feasibility Testing of Exergames for Patients with Mild Cognitive Impairment

Master's Thesis in Medical Engineering

submitted
by

Fabio Andre Egle

born 29.03.1995 in Mainz

Written at

Machine Learning and Data Analytics Lab
Department Artificial Intelligence in Biomedical Engineering
Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)

Advisors: Prof. Anne Koelewijn, Dr. Felix Kluge, Dr. Hartmut Lehfeld (Klinikum Nürnberg)

Started: 01.10.2020

Finished: 20.05.2021

Ich versichere, dass ich die Arbeit ohne fremde Hilfe und ohne Benutzung anderer als der angegebenen Quellen angefertigt habe und dass die Arbeit in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen hat und von dieser als Teil einer Prüfungsleistung angenommen wurde. Alle Ausführungen, die wörtlich oder sinngemäß übernommen wurden, sind als solche gekennzeichnet.

Die Richtlinien des Lehrstuhls für Bachelor- und Masterarbeiten habe ich gelesen und anerkannt, insbesondere die Regelung des Nutzungsrechts.

Erlangen, den 18. Juni 2021

Übersicht

Mild Cognitive Impairment (MCI) ist eine kognitive Beeinträchtigung, die sich in vielen Fällen zu einer Form von Demenz weiterentwickelt. Regelmäßiges körperliches und kognitives Training ist eine effektive Methode, um den Abbau der kognitiven Fähigkeiten zu verlangsamen. Kognitive Exergames auf der Basis von Inertial Measurement Units (IMU) und einem mobilen Endgerät können eine kostengünstige Möglichkeit sein, ein selbstständiges und motiviertes Training zu Hause zu ermöglichen. Diese Masterarbeit untersucht die Machbarkeit eines IMU-basierten Exergames für ältere Menschen mit MCI mit einem Fokus auf die Zuverlässigkeit der Bewegungsdetektion und die Benutzerfreundlichkeit des Systems. Zu diesem Zweck wurde ein Exergame als 3D Parkour in Unity mit einer rudimentären Bewegungsdetektion von 14 definierten Bewegungen entwickelt und in zwei Studien evaluiert. In einer initialen Bewertung der Bewegungsdetektion an aufgenommenen Bewegungsdaten von fünf älteren Menschen konnte das System mit einem F1-score von 77.05 % (macro-average) zeigen, dass es für einen Einsatz im Exergame ausreichend zuverlässig ist. Bei der Auswertung der Benutzerfreundlichkeit konnte das entwickelte Exergame Prototyp bei siebzehn jüngeren Menschen unter anderem exzellente Bewertungen von 86.47 in der System Usability Score erreichen. Auf diese Weise konnten in dieser Masterarbeit sowohl die Machbarkeit eines IMU-basierten Exergames, als auch die Bewegungsdetektion und Benutzerfreundlichkeit des entwickelten Protoypen demonstriert werden.

Abstract

Mild Cognitive Impairment (MCI) refers to a cognitive decline that in many cases progresses to a form of dementia. Regular physical and cognitive training is an effective way to slow down the decline of cognitive abilities. Cognitive exergames based on Inertial Measurement Units (IMU) and a mobile device can be a cost-effective way to provide independent and engaging home-based training. This master thesis investigates the feasibility of an IMU-based exergame for older people with MCI with a focus on the reliability of motion detection and the usability of the system. For this purpose, an exergame was developed as a 3D parkour in Unity with a rudimentary motion detection of 14 defined movements and evaluated in two studies. In an initial evaluation of the motion detection on recorded motion data of five elderly people, the system was able to show with an F1-score of 77.05 % (macro-average) that it is sufficiently reliable for the use in the exergame. In the evaluation of the usability, the developed exergame prototype was able to achieve an excellent rating of 86.47 in the system usability score among 17 younger people. As a result, this master's thesis was able to demonstrate both the feasibility of an IMU-based exergame as well as the motion detection and usability of the developed prototype.

Contents

Chapter 1

Introduction

According to the World Health Organisation, almost 10 million people are diagnosed with dementia every year and by 2050, the total number of people worldwide suffering from dementia is expected to reach 150 million [Wor19]. As the disease is a great burden for patients, relatives, and caregivers, it is particularly important to take action against the disease before it even evolves into dementia. In concrete terms, this means that the preliminary stage of dementia, mild cognitive impairment (MCI), must be addressed to prevent or slow down its progression [Pet14, Zha20].

Currently, there is only a very limited treatment for dementia and no therapy option for MCI [Pet16]. However, it has been shown that regular physical and cognitive training have a positive effect on the cognitive abilities of MCI patients [Lau19, Gat11, But18, Dem20]. One way to implement regular cognitive and physical training for MCI patients could be with specific cognitive exergames. With these movement-based games on a mobile device, television, or console, patients could counteract their cognitive decline with a high level of motivation and independence [Zha20]. However, a limitation of most current motion detection systems is that they are either not portable or expensive. The use of low-cost inertial measurement unit (IMU) sensors and a game on an existing mobile device could be a cost-effective alternative that can also be used by community-dwelling adults on a regular basis [O'R18].

While there are multiple studies which tested cognitive exergames with elderly or elderly with MCI, only a few used IMUs as input for the movement detection (Section 2.2). For this reason, this thesis will concentrate on the feasibility of an exergame for MCI patients which only relies on IMUs as input for the movement detection and assess the inherited reliability.

Another important factor for the usage of an IMU-based exergame is the usability and enjoyment of the system. An exergame is especially motivational due to its sensory feedback and independent usage at home. Both are factors contributing to a high adherence rate [Zha20].

Therefore, the enjoyment and usability of an exergame can be a first approach to evaluate an exergame and estimate usage behaviour.

In the context of this thesis, the reliability of movement detections with IMUs as well as the enjoyment and usabiliy are assessed on a MCI-Exergame prototype. For this purpose an MCI-Exergame prototype was developed, which connects to four IMUs and includes a basic movement detection of fourteen movements and walking on the spot. Concretely, the thesis aims to answer the following research questions with the help of the developed exergame prototype:

1. *Can different movements of old and young be reliably detected with inertial measurement units?*

2. *How usable is the developed exergame for elderly people with mild cognitive impairment?*

Two studies were conducted to answer these research questions. The movement recording study (Section 3.2) focused on establishing a development basis for the MCI exergame and tested different levels of movement instruction. The study included the recording of motion data from 5 elderly people with four IMUs capturing fourteen different movements as well as walking on the spot. In addition, the developed movement detection was tested on valid movements of this study. In a second study, the usability study (Section 3.3), the developed exergame was then tested regarding usability and enjoyment with seventeen younger participants. In addition, a first evaluation of the physical load and interaction with the system was conducted here.

The thesis is structured as follows. Chapter 2 - Fundamentals provides a brief overview on medical and technological basics that are important for understanding the thesis. Additionally, relevant papers on exergaming for MCI and movement detection are presented. Chapter 3 - Methods starts with the technical development of the MCI exergame prototype (Section 3.1). There, the developments regarding the data analysis of the raw sensor signals, the movement detection algorithms, as well as the gameplay and design are presented. Afterwards, details on both conducted studies are presented in separate parts. Chapter 4 - Results presents all outcomes of this thesis. Here, outcomes from the *MCI-Exergame* developement, the movement recording study, and the usability study are presented separately. In Chapter 5 - Discussion the main findings from the previous chapter are discussed, concluding in a final assessment of the feasibility and the answers to the two research questions regarding movement detection and usability. Finally, the thesis is concluded with Chapter 6 - Conclusion and Outlook, providing a final statement and a glance into the future.

Chapter 2

Fundamentals

The following chapter covers the medical fundamentals of mild cognitive impairment (MCI) and explains the usage of exergames in this context. Furthermore, the chapter presents technological and mathematical aspects that are important for the understanding of the thesis. This includes an overview of inertial measurement unit (IMU)s and the processing of their signals. Addionally, the chapter includes the presentation of individual studies on exergames for people with MCI and studies on motion detection with IMUs.

2.1 Mild Cognitive Impairment

Mild cognitive impairment (MCI) is generally defined by being a condition between age-appropriate cognitive abilities and a form of dementia [Pet99]. In contrast to dementia, MCI that has not yet led to any or only minor restrictions in the activities of daily living for the patient [Pet11]. Therefore, a MCI patient is able to mostly live independently and is not dependent on the help of others.

One decisive criterion for MCI for the diagnosis of MCI is that the patient is not suffering from any form of dementia yet. MCI is generally differentiated between two subtypes. With MCI, a distinction is generally made between amnestic mild cognitive impairment (aMCI) and non-amnestic mild cognitive impairment (naMCI). A patient with aMCI suffers from clinically significant memory impairment. In naMCI, attention, language skills, and visuospatial skills, among others, can be affected [Pet11]. Additionally, it is differentiated between a cognitive disorder in one domain and multiple domains (Figure 2.1) [B. 04, Pet04].

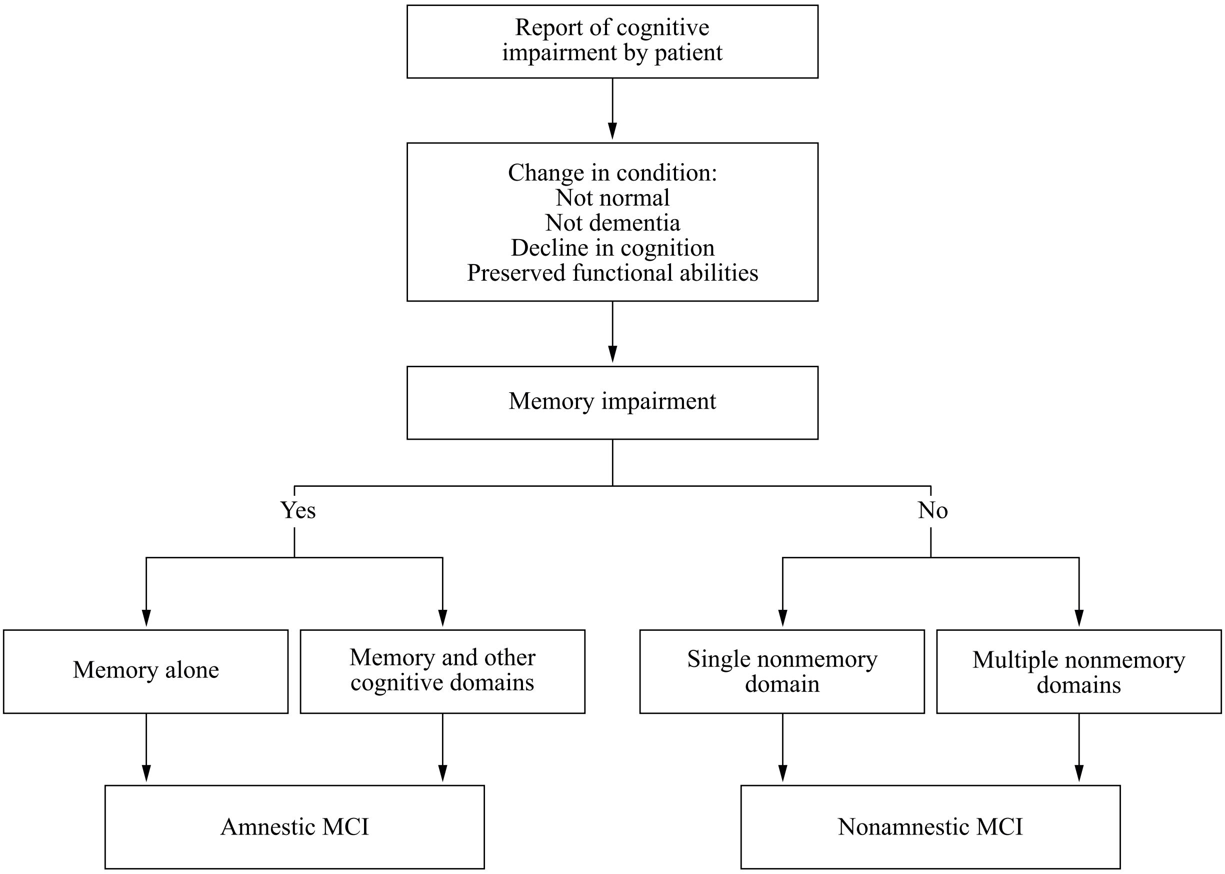


Figure 2.1: Criteria for MCI and subtypes [Pet11], adapted.

Not all cases of MCI are related to Alzheimer's disease (AD). While aMCI generally is connected to Alzheimer's disease, other etiologies can cause aMCI and naMCI as well. These include but are not limited to dementia with Lewy bodies, frontotemporal dementia, vascular cognitive impairment, and depression [Pet16]. Furthermore, other classifications, which are similar and complementary to the concept of MCI have emerged such as the pre-dementia phase neurocognitive disorder (NCD) from DSM-5 [Ame13] and prodromal AD [Dub07, Dub10, Dub14]. The origins and the differences between the terms in detail would lead too far in the context of this master thesis, therefore in this thesis only the term MCI is used. MCI is curently listed in the ICD-10-GM 2021 as F06.7 ”Leichte kognitive Störung“.

Many international studies have shown that MCI is a widely spread disease among elderly people. In these studies, the prevalence ranges from 12% to 18% for people over 60 years of age [Bus06, Di 07, Gan10, Lar02, Pet10]. In a newer meta-analysis by Gillis et al. [Gil19] from 2019 the MCI incidence per 1000 person-years was estimated 22.5 for an age of 75-79 years, 40.9 for an age of 80-84 years, and 60.1 for an age of 85 years and above.

People with MCI generally have a higher risk of developing dementia. In a study by Roberts et al., 28.7% of the participants with a form of MCI progressed to a form of dementia in a median follow-up of 5.1 years. [Rob14]. In a meta-analysis by Mitchel and Shiri-Feshki [Mit09], the annual conversion rate from MCI to dementia, Alzheimer's disease, and vascular dementia was 9.6%, 8.1%, 1.9%, respectively, in a specialist clinical setting and 4.9%, 6.8% and 1.6% in community studies.

Despite the high prevalence in elderly patients and many studies investigating possible treatments, there are currently no approved pharmacologic treatments for MCI. Due to this reason, non-pharmacologic treatments like regular physical activity and cognitive training have been thoroughly studied in the last years. It has been clearly demonstrated that regular exercise has positive effects on the cognitive abilities of MCI patients [Lau10, Lau19]. Similarly, positive effects on the patient's cognitive abilities were observed from cognitive training [Gat11, But18]. In a recent review by Demurtas et al., it was concluded that physical activity has a positive effect on several cognitive outcomes in MCI and dementia patients. As the findings are only supported by low-to-moderate certainty, the authors argued that and more better-structured randomized controlled trials (RCT) with higher sample sizes are needed to confirm this indication [Dem20]. Additionally, it is indicated that a combination of physical activity and cognitive training could not only increase the cognitive training effect in MCI patients but also have a moderate-to-large positive effect on activities of daily living (ADL) [Kar17].

In this regard, serious games with cognitive elements, especially exergames with physical and cognitive elements, could play a significant role towards the objective of implementing cognitive training in the daily life of MCI patients.

2.2 Exergames for Mild Cognitive Impairment

Exergames can be defined as "digital games that require bodily movements to play, stimulating an active gaming experience to function as a form of physical activity" [Ben18]. To be able to interact with the game, the system relies on integrated or external sensor technology depending on the gaming device. Exergames have been developed and tested with different technological implementations. These include among others: motion capture systems, inertial sensor systems, depth-sensing with structured light, virtual reality systems, tap mats, balance boards, and ergometers [O'R18, RO17, GA, AH18, Chu21]. Depending on the realization, exergames can be played on smart TVs, computers, tablets, mobile phones, and gaming consoles [Adc19, AH18, Sch13, Amj19]

The effect of cognitive and physical exergames on MCI patients was already the subject of several studies in the last years. In the following, a selection of relevant studies is presented.

In a RCT by Amjad et al. [Amj19], 44 subjects with MCI were either performing cognitive games on an Xbox 360 Kinect or doing motion exercises in an active control group for a duration of 6 weeks. The games were selected from the game *Body and Brain Exercises by Dr Kawashima* in the categories logic, physical, memory, reflexes and math. The effect was measured through electroencephalogram (EEG) and cognitive tests like the Montreal cognitive assessment (MoCA) and trail-making-test (TMT) A & B. After 6 weeks of training cognitive tests improved significantly in the exercise group, whereas the effects in the control group were minimal. Additionally slowness and complexity of thee EEG parameters improved significantly, suggesting a positive impact on the progression of the cognitive decline.

In another study by Bamidis et al. [Bam15] used computerized cognitive and physical games and training to investigate the effect on global cognition, which was calculated by the mean individual performances in working memory, episodic memory, and executive functions. The group of 322 community-dwelling older adults consisted of individuals who ranged from cognitive healthy subjects to people with MCI and dementia. While a significant general change in global cognition was found compared to the control group, it was indicated that intervention effects on global cognition were lower in people with MCI than healthy subjects and respectively lower in participants with dementia than participants with MCI. Bamidus et al. conclude that this finding supports the strategy that prevention of neurocognitive disease must begin long before ADL impairment.

In A Cluster Randomized Clinical Trial by Anderson-Hanley et al. [AH12], 63 participants completed a 3 month cybercycling or traditional exercise plan. Both groups rode identical stationary bikes, but just the cybercycling group had an additional display enabled. The participants could experience 3D tours, compete with their last best ride, and try to outpace it. Cognitive Outcomes were measured via Color Trails 2-1, Stroop C and Digit Span Backwards tests. Anderson-Hanley et al. found an increased effect on executive functions in the cybercycling group over the traditional exercise group. Additionally, a significant rise in brain derived neurotrophic factor (BDNF) was found and fewer participants converted to MCI (3 instead of 9) in the cybercycling group. Therefore, the authors argued that cybercycling could have a preventive effect on the participants.

In an RCT by Adcock et al., [Adc19] 37 healthy adults, 65 years and older played home-based exergames, which consisted of Tai Chi-inspired exercises, dancing and step-based cognitive games. The system "Active@Home" consisted out of four inertial measurement units and an HDMI dongle to insert into a monitor. MRT-imaging brain volume was assessed, which did not exhibit a significant change or difference to the control group. Furthermore, inhibition and working memory significantly increased in the intervention group compared to the control group.

In a systematic review by Zhao et al. [Zha20], it was concluded that exergames can improve cognitive and physical functions to some extent, but that further high-quality studies are needed. Zhao et al. argued that due to different equipment, set-ups, training frequencies, and durations, interventions are difficult to compare and it makes it difficult to determine which exergame type brings the greatest cognitive benefit. Additionally, the authors state that exergames are useful options for elderly people with MCI as they can be inexpensive, safe, and allow usage in an unsupervised environment like the patient's home. Furthermore, the sensory feedback and stimulation of the exergames can provide a motivational factor and can achieve a high adherence rate [Zha20].

In a systematic review and meta-analysis by Wollesen et al., a preliminary minimal dose of 60 min per week was recommended [Wol20]. For interventions away from home, this would mean that community-dwelling elderly need to regularly travel distances to specific facilities, which would be an additional obstacle and a time-consuming and potentially costly matter. A home-based, low-cost exergame, which could be played on the own TV, tablet or computer, could be a viable solution to overcome these problems and allow for regular cognitive training and exercise. One efficient solution could be the usage of a combination of low-cost inertial measurement unit (IMU) and already existing devices like a smartphone/tablet and a TV.

The following outcomes from these studies are particularly important to the development of this master's thesis:

- Exergames can provide a cognitive benefit for elderly people and MCI patients.
- Through the playful environment of exergaming, it is possible to achieve a higher motivation in the patient than through conventional sports methods or cognitive training.
- Simultaneous cognitive and physical training can have a particularly large impact on cognitive ability and ADLs.
- Exergames, with inexpensive equipment that can be played at home without supervision, represent a simplified access to this training opportunity.
- Inexpensive IMUs could be a viable option for this purpose.

2.3 Inertial Measurement Unit

An inertial measurement unit (IMU) is a wearable device that generally consists of accelerometers and gyroscopes and allows the measurement of 3D acceleration and 3D angular velocity. In the next step, these values can be used to calculate parameters such as orientation, velocities and trajectories in space. As these sensors are very light-weighted, easily attachable and allow precise measurements, they are already part of many day-to-day, scientific and industrial applications. These include inertial navigation [Ben15], biomedical applications such as gait [Klu17] or biomechanical motion analysis [Dor20], and smartphone apps for playing or sports[Adc19].

Depending on the use case of the sensor, the composition of an IMU differs. Therefore, an IMU can additionally incorporate a magnetometer or barometer, among others.

In 1852, Jean Bernard Leon Foucault first demonstrated a pendulum to measure earth rotation. Over the years, gyroscopes were adapted to be used as a stabilization and navigation instrument in ships and aircrafts. First, strain gauge accelerometers were developed in the 1920s and were large, heavy, and expensive. Nowadays, accelerometers and gyroscopes are micro-electro-mechanical systems (MEMS), which means that they measure the mechanical changes in the device and process them electronically. [Ben15].

Accelerometer and gyroscope in an IMU measure their respective values in 3 axes (Equation 2.1). It can be useful to form a right-handed coordinate system out of three axes from the sensor perspective, the so-called sensor frame S. As the sensor is moving and changing orientation, the sensor frame S is changing with it.

$$\overrightarrow{a_S} = \begin{pmatrix} a_{S,x} \\ a_{S,y} \\ a_{S,z} \end{pmatrix}; \quad \overrightarrow{\omega_S} = \begin{pmatrix} \omega_{S,x} \\ \omega_{S,y} \\ \omega_{S,z} \end{pmatrix}; \tag{2.1}$$

On the other hand, the world frame W is the fixed frame and serves as a reference for the representation of other frames such as the sensor frame S. In this respect, the orientations of the sensor frame S are always in reference to the world frame W (Equation 2.2). In the context of this thesis, the world frame W is defined so that gravity is negatively aligned with the axis W_z. The heading value in turn depends on the initial orientation as no magnetometer was used for further specification. Therefore, in this thesis, the axis W_y is in the assumed initial line of sight of the subject. The orientation of S in reference to W can be described by the quaternion q_S. For most part of this thesis the body frame B is used instead of S. The body frame is additionally takes the orientation of the sensor in the IMU case into account.

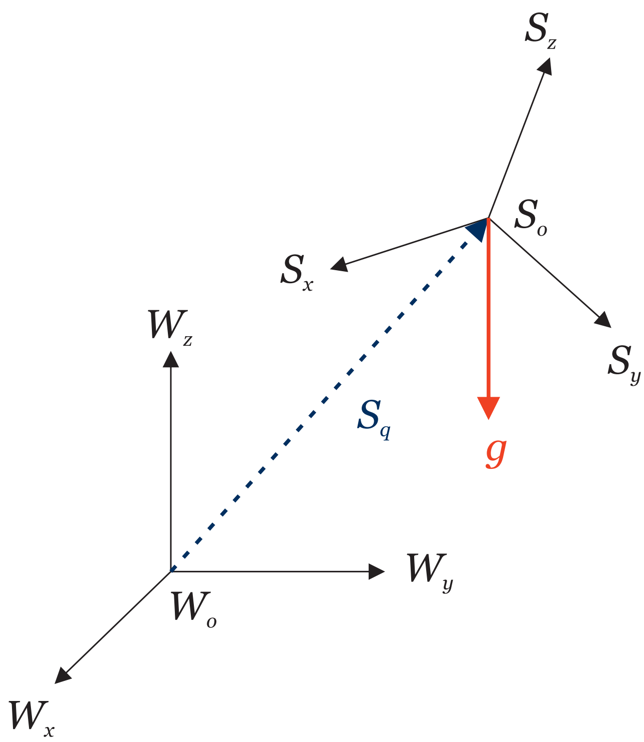


Figure 2.2: World frame and sensor frame. Own representation based on [Kui99]

In a stationary setting, the accelerometer only measures the gravitational force, thereby providing information about the direction of gravity in the sensor frame S and the orientation of the sensor in world frame W. As soon as the accelerometer setting is dynamic, the gravity effect is combined with the accelerating motion to a joined value, posing a problem for pure acceleration, velocity, and trajectory calculations. If the orientation of the inertial sensor in the world frame is known, gravity can again be subtracted after calculating a_W from a_S. By subtracting gravity g_W from a_W, gravity-removed acceleration $a_{W,grm}$ is calculated.

$$\overrightarrow{a}_{W,grm} = \overrightarrow{a}_W - \overrightarrow{g} \tag{2.2}$$

Given an orientation, e.g., in the quaternion form q_S of the sensor frame S in reference to the world frame W the acceleration vector can be described in the world frame W. From here, after the removal of gravity, $a_{W,grm}$ can theoretically be integrated for velocity and double-integrated to obtain the trajectory of the IMU. The current orientation for this procedure can be obtained in different ways. In static conditions, it would be possible to use the accelerometer's gravity measurement to obtain the orientation without heading angle. In turn, in dynamic conditions, given an initial orientation, the gyroscope's angular velocity could be used to get information about the current orientation. Due to noise and small measurement errors, integrating the angular velocity results in an increasing orientation error [Lui05]. There have been several sensor fusion approaches to account for this problem. These sensor fusion algorithms combine signals from the accelerometer, the gyroscope, and, if applicable, the magnetometer. Among these are the Kalman filter [Kal60] and complementary filter (CF) like the Madgwick filter [Mad11]. Since the Madgwick filter is used in the practical development of this thesis, it will be briefly explained in the following.

The Madgwick Filter by Madgwick et al. [Mad11] is an orientation estimation algorithm in quaternion representation combining an accelerometer $\overrightarrow{a}_B$ (q_{a_B} as pure quaternion), gyroscope $\overrightarrow{\omega}_B$, and (if applicable) magnetometer data and thereby compensating gyroscope drift. The main idea is here to calculate the next orientation ${}^B_W q_{est}$ by the combination of the integration of gyroscopic quaternion ${}^B_W \dot{q}_\omega$ and the orientation obtained by the gravity measurement $\overrightarrow{g}_W$ (q_{g_W} as pure quaternion) by the accelerometer. This way, the accelerometer measurements are used as a long-term correction of the gyroscopic drift. While the angular velocity measured by the gyroscope is integrated, an optimization problem is formulated for the accelerometer measures to minimize the difference between the gravity transformed in the body frame and the measured accelerometer values.

$$\min f\left({}^B_W q, q_{g_W}, q_{a_B}\right) \tag{2.3}$$

$$f\left({}^B_W q, q_{g_W}, q_{a_B}\right) = {}^B_W q^* \otimes q_{g_W} \otimes {}^B_W q - q_{a_B} \tag{2.4}$$

To solve the optimization problem, the Madgwick filter uses one iteration of a gradient descent algorithm for each sample yielding the objective function loss ∇f. The orientation is then updated by fusing both information.

$${}^B_W \dot{q}_{est,t} = {}^B_W \dot{q}_{\omega,t} - \beta \frac{\nabla f}{\|\nabla f\|} \tag{2.5}$$

$${}^B_W q_{est,t} = {}^B_W q_{est,t-1} + {}^B_W \dot{q}_{est,t} \Delta t \tag{2.6}$$

The filter is characterised by the fact that it has a low computational load at high accuracy and delivers reliable results at low sampling rates [Mad11]. Due to these features, the Madgwick algorithm was assessed particularly suitable for use in real-time motion detection in the context of this thesis. As the Madgwick algorithm uses gravity as a correctional factor, it cannot account for the gyroscopic drift around the global axis without the incorporation of a magnetometer.

2.4 Movement Detection with Inertial Measurement Units

IMU-based systems can be considerably cheaper then other movement detection or pose estimation systems like motion capture or depth-camera. Additionally, they provide unique features as they are more independent from external factors such as location or lightning. Consequently, there has been a considerable focus studies on movement detection through IMUs. [O'R18] There have been many different approaches to movement detection and pose estimation with IMUs in the last years. The following is a selection of relevant approaches in the field of movement detection. Basically, a first distinction can be made between the usage of machine learning methods and approaches without machine learning. Furthermore, the approaches differ greatly depending on the body region, the number of IMUs used, the number and specificity of movements to be classified, and real-time capability.

Wu et al. developed a gesture recognition system based on joint movements, joint angles, and arm orientation. The system used three IMUs, each equipped with a 3-axis accelerometer, gyroscope, and magnetometer. One IMU was fastened on the lower right arm, on the upper right arm, and on the upper right body side. The raw accelerometer values were fused by a Kalman filter into an orientation in quaternion form, which then was streamed to a computer via Bluetooth. For the gesture classification, a Movement Likelihood Matrix was used and the gesture with the highest similarity was detected. This way, the system was capable of recognizing twelve different gestures with an average accuracy of 91.86%.

Rajkumar et al. [Raj21] developed and evaluated a exergame system with IMUs. The system was based on five 9DOF IMUs, fixated on both upper and lower arms as well as the lower back. While there was no direct movement classification involved, the system was able to portrait the upper limp movements of the participants in real-time by calculating the relative orientation between the IMUs. This was then compared with the Kinect version of the pose estimation, where Bland-Altman limits of agreement showed ±10° for the exercises in the coronal and transverse planes, suggesting that the system is able of reliably visualising the players movements.

As turning can be a major difficulty for patients with movement disorders, Mahmoud El-Gohary et al. [EG13] developed an algorithm to detect and characterize turning movements. This setup consisted of one IMU with 3-axis accelerometer, gyroscope, and magnetometer, which was worn on the lumbar spine. In essence, the angular velocity was transformed to the world frame W with the IMU-provided quarternion. Afterwards, the relative turn angle was calculated with the z-axis component of the angular velocity, which was summed up depending on the duration and direction. The developed algorithm detected movements with a maximum sensitivity of 0.90 and a maximum specificity of 0.75, depending on the evaluation method.

In a recent publication by Hua et al. [Hua20] machine learning methods to classify nine upper extremity exercises were evaluated. These movements included *standing row, external rotation with abducted arm, external rotation, biceps curl, forearm pronation, wrist curls, lateral arm raise, front arm raise and horizontal abduction* Participants were equipped with three IMUs on the right arm and one on the lateral torso. With calculated joint angle data from IMU orientations Random Forest, Linear SVC, k-nearest Neighbours, and Multilayer Perceptron were used to differentiate between movements. Random forest models with flattened kinematic data performed best with an accuracy of 98.6%.

Drumond et al. [Dru18] used an LSTM based approach to classify five different actions, including *walk*, *idle*, *run*, *swing* and *crouch*. Two IMUs were placed on each lower arm and data was collected from 11 participants at a frame rate of 50 Hz. The neural network was trained on 20 features, which consist the quaternion orientation and acceleration values for each IMU. For a window size of 60 samples, the neural network yielded an accuracy of 96%.

2.5 Related Patents

The following list contains patents related to cognitive exergaming (1.), portable exergaming (2.), interactive fitness training (3.), and IMU-based video game input (4.). More information on each item can be found in Appendix A.

1. **Interactive physical & cognitive exercise system and method - US20160293033A1**

2. **Systems and methods for portable exergaming - US10039981B2**

3. **Method and system for interactive fitness training program - US8113991B2**

4. **Methods and apparatus for a video game magic system - US9545571B2**

Chapter 3

Methods

This thesis investigates the feasibility of exergames for persons with mild cognitive impairement with a focus on motion detection and usability. For this purpose an *MCI-Exergame* prototype was developed and two studys were conducted. In the first study movement data of elderly people was recorded and evaluated. The second study served the purpose of evaluating different factors of the usability on the prototype exergame. The methods for each of the stated components of the thesis are described in a separate section in this chapter.

3.1 Exergame Development in Unity

The following section describes the development steps on the way to the *MCI-Exergame* prototype. First, the technology employed in the project is described. Then the processing of the data and the calculation of the parameters for the movement detection are presented. After the explanation of the detection algorithms that were used, the last part focuses on the development of the gameplay and the associated design choices.

Since the capabilities of the movement detection, game logic and design highly depend on each other, the development process was conducted in a parallel workflow. The objective was to develop an exergame that provides continuous aerobic movement and challenges the elderly people sufficiently, both cognitively and physically, through the interaction with various 3D objects. Also, it should characterised by a high level of usability and enjoyment. Another goal was that the game should be as self-explanatory as possible and allow independent use at home. It was decided to implement this game as a virtual parkour in which the player can interact with different objects by performing the matching movement.

3.1.1 Technology

For the development of this thesis, several different hardware and software solutions were used. These included, among others. inertial measurement units (IMU), the Unity development environment and the 3D design application Blender.

Four NilsPods V1 IMU were provided by the Portabiles GmbH [por21]. A NilsPod V1 includes a 3-axes accelerometer and gyroscope as well as a barometer and is capable of Bluetooth low energy (BLE). Through BLE, values can be streamed in different preset frequencies. The accelerometer and gyroscope have an adjustable range of $\pm 2 - \pm 16g$ and $\pm 125 - \pm 2000°/s$ respectively. Due to its small size, reliable wireless connectivity, long battery life and already tested use in the scientific environment [Cer20, Aho19], the NilsPod is particularly well suited for scientific work in motion analysis.

To develop the *MCI-Exergame* prototype, the development environment Unity [Uni20] was chosen. Unity can be used for the development of two-dimensional, three-dimensional, virtual reality and augmented reality games and simulations. Unity has even been used before for a cognitive exergame development [Adc19]. It is a cross-platform game engine that uses C# as the main scripting language. The Unity update cycles allow users to manage game objects and to construct scenes by assigning object properties and executing code scripts over time. In addition to the capability to design 3D objects and user interface elements in Unity directly, Unity supports 3D objects and animations created in other 3D design and CAD programs. One of these is the commonly used open-source application Blender [Ble21], which includes features like 3D modelling, animation building and the rigging of characters. For the developments in this thesis, Unity Versions 2020.1.8f1 - 2020.3.3f1 and Blender V2.91 were used. Furthermore, assets from the Unity Asset Store or external compatible packages can be added to Unity. The following external resources were implemented and used in the *MCI-Exergame* prototype:

- **Unity Asset Store:** Unity Standard Assets [Uni20], Bluetooth LE for iOS tvOS and Android [Sha20], Simple JSON, Fantasy Skybox FREE [Ren20], FREE Stylized PBR Textures Pack [Lum18], Better Streaming Assets [Gwi19]
- **NuGet Packages:** MathNet.Numerics [Ruea], MathNet.Filtering [Rueb]
- **Other Sources:** Madgwick C# implementation [x-i21]

The Math.Net packages were used to implement matrices calculation as well as IIR filters. The Bluetooth LE asset enabled the developement of an convenient communication with the NilsPods via BLE, which works on macOS, iOS and Android. The Better Streaming Assets were needed to implement the calibrations as retrievable files in a sustained storage even after the application is compiled for native usage in iOS and Android. The Simple JSON enabled an easy load of JSON files (like the calibration). While all 3D models were developed by the author, textures and materials were partly obtained from Unity Standard Assets, Fanstasy Skybox FREE and FREE Stylized PBR Textures Pack. The gameplay developemnt with Unity and Blender is further explained in Section 3.1.4.

3.1.2 IMU Data Processing

The following part of the thesis presents the procedure from incoming IMU data until the calculation of parameters like velocities and positions. While the movement recording study (Section 3.2) was carried out with a sampling frequency $f_{S1} = 102.4Hz$, for the game development $f_{S2} = 51.2Hz$ was chosen for performance and compatibility reasons. All features in the game are implemented for both sampling rates. Each package received via BLE can include one or more samples. For this reason, the package is first split and the data for each sample is processed separately (Figure 3.1). The processing steps for the IMU Data consist of calibration, orientation estimation and velocity and trajectory calculations. As the aim was to develop a system that not only processes data in real time, but is also capable of directly displaying the movements of the sensors, this initial data analysis is performed directly for each sample that is streamed. Afterwards all parameter needed for the movement detection are queued in a buffer, which includes the data of the last 3 seconds (3 f_s samples). The following list contains all the parameters that the movement detection requires from the initial data analysis and that are stored in a buffer.

1. Local acceleration $\overrightarrow{a_B}$ and local angular velocity $\overrightarrow{\omega_B}$
2. Global gravity-removed acceleration $\overrightarrow{a}_{W,grm}$ and global angular velocity $\overrightarrow{\omega}_W$
3. Orientation ${}^B_W q$ of the sensor (body frame) in the World frame W in quaternion form
4. Barometer altitude x_p and the current difference to the reference barometer h_{diff}
5. Global velocity vector $\overrightarrow{\dot{x}}$ and position vector $\overrightarrow{x}$ of the IMU
6. Reference orientation for *turning* q_{ref} (leg-placed IMUs only)

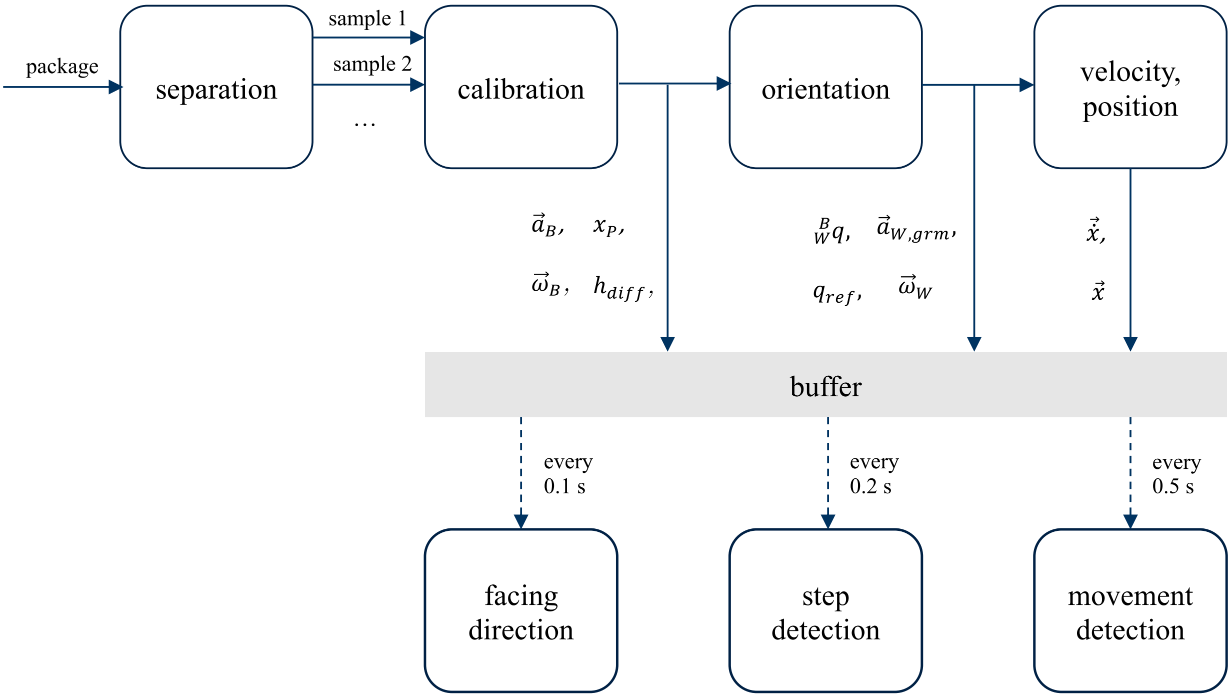


Figure 3.1: Flowchart presenting the the individual steps from package reception to movement detection.

Calibration

The calibration aims to compensate for the misalignment between the sensor and body frame. As the body frame is dependent on the IMU-case, the calibration also adapts the measurements to the properties of the individual IMU. The Ferraris calibration [Fer] is an effortless in-field calibration method, which requires measurements from six static orientations and three rotations around the three IMU axes. The Ferraris calibration estimates the following parameters: The biases in gyroscope b_g and accelerometer b_a, the scale factors K_g and K_a, the gyro sensitivity due to acceleration $k_{g,a}$ and the orientation matrices R_g and R_a. From these factors and the measured values $\vec{u}_a$ and $\vec{u}_g$, $\vec{a}_B$ and $\vec{\omega}_B$ can then be determined.

$$\vec{a}_B = R_a^{-1} \cdot K_a^{-1} \cdot (\vec{u}_a - \vec{b}_a) \tag{3.1}$$

$$\vec{\omega}_B = R_g^{-1} \cdot K_g^{-1} \cdot (\vec{u}_g - \vec{d}_g - \vec{b}_g); \quad d_g = K_{g,a} \cdot \vec{a}_B \tag{3.2}$$

The required calibration coefficients for the NilsPods were provided by Portabiles GmbH [por21] as JSON files. The coefficients for the four pods and many more are stored in the

MCI-Exergame prototype and are loaded accordingly as soon as a NilsPod is selected in the *Menu*. The required matrix calculations were implemented using the MathNet.Numerics package [Ruea].

Sensor placement and orientation

The development of the MCI exergame was designed to use four IMUs with predetermined placements and orientations. Two IMUs are to be fastened on the upper side of the wrist (distal radioulnar joints) with a bracelet. The orientation of the NilsPods on the wrists was correct if the lower end of the "P" was pointing towards the wrist. The other two IMUs are fixed with a clip on the shoe at the back of the foot (dorsal foot) (Figure 3.2). The orientation of these was correct if the lower end of the "P" was pointing towards the tip of the foot.

(a) Placement on the right wrist (b) Placement on the left shoe

Figure 3.2: Pictures of the placements of the NilsPods.

The right-handed coordinate system of the NilsPod V1 negatively aligned with gravity in the direction of the z-axis B_z when laid flat (Figure 3.3). As the NilsPod V1 does not incorporate a magnetometer and thereby cannot determine the initial heading values dependent on the earth magnet poles, the initial orientation was externally set. For this purpose, the initial line of sight of a person during the connection phase was used as the global y-axis W_y. The initial orientation of each NilsPod depending on the placement was then assumed and set to an initial orientation in quaternion form. The initial orientation were set to the following quaternions (Euler angles are given for better visualization):

- **Right arm:**

  Quaternion: $\left[x : 0.27059805 \quad y : -0.65328148 \quad z : -0.27059805 \quad w : -0.65328148\right]$

  Euler angles: $\left[X : -90^\circ \quad Y : 45^\circ \quad Z : -90^\circ\right]$

- **Left arm:**

  Quaternion: $\left[x : 0.65328148 \quad y : -0.27059805 \quad z : -0.65328148 \quad w : -0.27059805\right]$

  Euler angles: $\left[X : -90^\circ \quad Y : -45^\circ \quad Z : -90^\circ\right]$

- **Legs:**

  Quaternion: $\left[x : 0.65328148 \quad y : -0.27059805 \quad z : -0.27059805 \quad w : -0.65328148\right]$

  Euler angles: $\left[X : -45^\circ \quad Y : -0^\circ \quad Z : -90^\circ\right]$

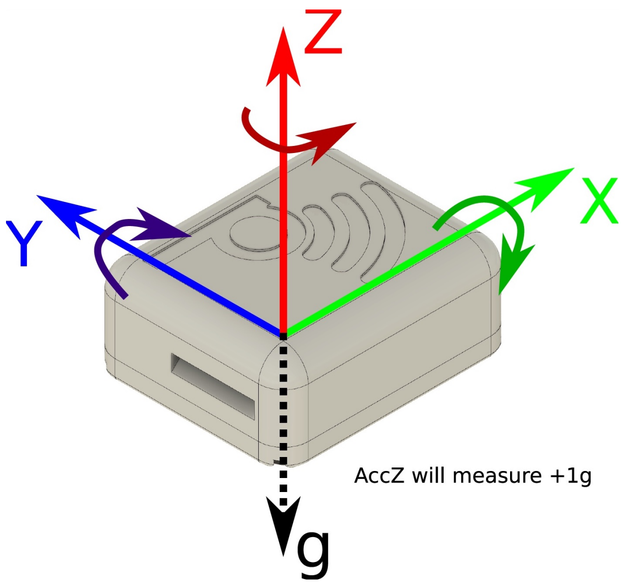


Figure 3.3: Shows local coordinate system of the NilsPod V1. When laid flat gravity negatively aligns with local z-axis [por21].

The initial orientation of each NilsPod was then adjusted to each processed sample after calibration. For this purpose the Madgwick [Mad11] complementary filter was used. Given an initial orientation, the Madgwick filter processes the accelerometer and gyroscope values yielding an orientation in quaternion form. A brief overview on the Madgwick algorithm can be found in Section 2.3. The Madgwick filter implementation in C# from x-io Technologies Limited [x-i21] was used. As the set initial orientation is naturally different from the real orientation with respect to the gravity direction, the Madgwick filter gain β was set to high value of $\beta = 3$ for the first second. The parameter β is a tunable parameter that represents the gyroscopic error in the direction of accelerometer measurements [Nit19]. A higher value strengthens the impact of the accelerometer and changes the orientation to be aligned with the gravity vector. During this time, no velocities or trajectories are calculated. After the first second, the parameter changes to the value of $\beta = 0.046$. As the Madgwick algorithm uses the gravity as reference against to compensate the drift of the gyroscope, the heading angle around the cannot be adjusted. It is all the more important that the heading of the intial orientation is approximately correct. The execution of a movement or pose to detect the heading angle could be a useful addition in this context.

Velocity and position

After the orientation ${}^{B}_{W}q$ is obtained, $\overrightarrow{a_B}$ and $\overrightarrow{\omega_B}$ are transformed to the World frame W and gravity is removed from the acceleration (Equations 3.3-3.4, 2.2). To transform vectors by a quaternion, the vector needs to be transcribed into a quaternion by setting x, y, z values as q_x, q_y, q_z respectively and adding $q_w = 0$ [Kui99].

$$q_{\omega_W} = {}^{B}_{W}q \otimes q_{\omega_S} \otimes {}^{B}_{W}q^{-1} \tag{3.3}$$

$$q_{a_W} = {}^{B}_{W}q \otimes q_{a_S} \otimes {}^{B}_{W}q^{-1} \tag{3.4}$$

After gravity removal, the z-axis component $a_{W,grm,z}$ is filtered with a 3rd order high pass Butterworth filter [But30] with a cutoff frequency of $0.02Hz$ to eliminate any remaining offset which could cause the velocity to drift over time. Then, the global acceleration is processed by an 8-point centered moving average. For the implementation of the Butterworth filter [But30], the MathNet.Filtering [Rueb] package was used.

For the calculation of velocity and position, the global z-axis component was again processed separately. To calculate the vertical velocity and position a complementary filter of acceleration and barometer altitude was implemented [Sab14]. Sabatini and Genovese proposed to use an extended kalman filter to estimate the orientation and transform acceleration and then remove the gravity component. Afterwards they used their developed complementary filter to fuse vertical acceleration and barometer altitude to obtain vertical velocity and position. Instead of the extended alman filter used by Sabatini et al., the already described Madgwick implementation was used to yield the orientation quaternion for the rest of the algorithm. In the sensor fusion step (Equation

3.5) the last velocity $\dot{x}_{z,k-1}$, position $x_{z,k}$, the sample time T_S, the complementary filter gain K_c, the difference in position $\Delta x_{z,k}$ and the change in velocity $\Delta v_{z,k}$ are used to calculate the current velocity and position:

$$\begin{bmatrix} x_{z,k} \\ \dot{x}_{z,k} \end{bmatrix} = \begin{bmatrix} 1 & T_s \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{z,k-1} \\ \dot{x}_{z,k-1} \end{bmatrix} + \begin{bmatrix} 1 & T_s/2 \\ 0 & 1 \end{bmatrix} K_c \cdot T_s \Delta x_{z,k-1} + \begin{bmatrix} T_s/2 \\ 1 \end{bmatrix} \Delta v_{z,k-1} \quad (3.5)$$

Whereas $\Delta x_{z,k}$ is defined as $\Delta x_{z,k} = x_{p,k} - x_{z,k}$ and $\Delta v_{z,k}$ as $\Delta v_{z,k} = T_S \ a_{z,k}$.

Furthermore, the complementary filter gain K_c is based on the standard deviations of the noise of acceleration σ_w and altitude σ_v:

$$\mathbf{K}_c = -\begin{bmatrix} \sqrt{2\sigma_w/\sigma_v} \\ \sigma_w/\sigma_v \end{bmatrix} \quad (3.6)$$

For this thesis, the ratio of $\sigma_w/\sigma_v = 1/2$ from the method A of the original paper [Sab14] was adopted. This ratio empirically showed the best properties for a fusion between barometer and accelerometer.

Barometer altitude h was calculated by converting the measured pressure with respect to the standard pressure level $p_0 = 1013.25hPa$.

$$h = 44300 \left(1 - \left(\frac{p}{p_o}\right)^{0.19}\right) \quad (3.7)$$

In contrast to the study by Sabatini and Genovese, the difference between the two barometers was calculated in the exergame to better compensate pressure changes in the environment, before the 8-point centered moving average was applied.

$$x_p = h - h_{ref} \quad (3.8)$$

The value of the reference IMU was only updated when the length of the Euclidean norm of the vector $a_{w,grm}$ was below a value of $0.2m/s^2$ to not update the reference during movements. Additionally, during these samples, the absolute difference between barometers was adapted to map the pod visualization (Section 3.6) to the origin with a lerping factor $l_h = 10/T_s$. Initially, h_{diff} is set to the mean difference of both pods in the first two seconds of streaming.

$$h_{diff,k} = h_{diff,k-1} + (x_p - h_{diff,k-1})/l_h \quad (3.9)$$

The calculation of the positions in the global XY-plane is only for visualisation purposes and has no part in the movement detection. To calculate $v_{k,x}$, $v_{k,y}$ the accelerations $a_{k,x}$ and $a_{k,y}$ are multiplied by the sampling time and added to the last velocity.

$$\overrightarrow{v}_{k,xy} = \overrightarrow{v}_{k-1,xy} + \overrightarrow{a}_{k,xy}\, T_S \tag{3.10}$$

To dedrift the resulting velocity, a 2nd order high pass Butterworth [But30] with a cutoff frequency of 0.1 Hz was used. The position $\overrightarrow{x}_{xy}$ was updated depending on the velocity magnitude and otherwise shifted back to the origin with $l_O = 7/T_s$

$$\overrightarrow{x}_{xy,k} = \begin{cases} \overrightarrow{x}_{xy,k-1} + \overrightarrow{\dot{x}}_{xy,k-1}\, T_S & if\ |\overrightarrow{\dot{x}}_{k-1}| > 0.5 \\ \overrightarrow{x}_{xy,k-1} - \frac{\overrightarrow{x}_{xy,k-1}}{l_O} & else \end{cases} \tag{3.11}$$

To be able to detect rotations in the motion detection, a reference orientation q_{ref} for the NilsPods on the legs was also calculated and continuously updated to compensate for the drift in heading angle. This orientation in quaternion form is first defined with the current quaternion and then linearly interpolated with each new orientation with a factor $l_r o = 60/T_S$.

All parameters for the movement detection mentioned above are buffered for a window of three seconds.

3.1.3 Movement Detection

This section describes the detection algorithms used in this thesis. Similar to the gameplay of the MCI exergame, the movement detection is an experimental prototype. Many of the aspects of the following algorithms need to be further developed, adapted, optimized and validated. Addionally, the aim of the movement detection was, that not every remotely close movement would be detected, but that a specific quality of movement would be archieved. The game should tell the player which movement should be performed and only detect it if that movement was correctly done. All of the following algorithms are based on evaluating the current parameter buffer (Table 3.1) of $3f_s$ samples. For all detections except step detection, the window is evaluated every 0.5 seconds. For a smoother game experience, step detection is evaluated every 0.2 seconds. In addition, the player's orientation is evaluated every 0.1 seconds. Since the repeated routines depend on the Unity cycle and are thereby frame-dependent, it is not always guaranteed that they will be executed at the exact time. For all detections algorithms but the *step detection*, the *facing direction* and the *step over*, the window is first trimmed from both sides until the magnitude of the the gravity-removed acceleration $a_{W,grm}$ is higher or equal than $0.3\ m/s^2$. This trims the window

movements	$\overrightarrow{a_B}$	$\overrightarrow{\omega_B}$	$\overrightarrow{a}_{W,grm}$	${}^B_W q$	h_{diff}	$\dot{\overrightarrow{x}}$	$\overrightarrow{x}$	q_{ref}
reaching up			RA/LA	RA/LA	RA/LA		RA/LA	
picking up			RA/LA	RA/LA	RA/LA		RA/LA	
boxing	RA/LA		RA/LA					
clapping	RA&LA		RA&LA	RA&LA				
jump move		RA&LA	RA&LA			RA&LA		
stepping over		RL/LL	RL/LL				RL/LL	
turning				RL/LL				RL/LL
stepping sideways			RL/LL			RL/LL		
step detection		RL&LL						

Table 3.1: Overview which parameters from which NilsPod are used in which movement detection. RA: right arm, LA: left arm, RL: right leg, LL: left leg, $\overrightarrow{a_B}$: local acceleration, $\overrightarrow{\omega_B}$: local angular velocity, $\overrightarrow{a}_{W,grm}$: global gravity-removed acceleration, ${}^B_W q$: orientation of the sensor, h_{diff}: difference to the reference barometer, $\dot{\overrightarrow{x}}$: global velocity, $\overrightarrow{x}$: position, q_{ref}: reference orientation.

to the part where there actually is movement and when saving a detection, a smaller window can be specified where the movement occurred.

Step detection

The step detection is based on local gyroscope values. For this detection, it is assumed that a fast negative angular velocity around the local y-axis B_y is measured in the leg-based IMUs when lifted and a positive angular velocity when the foot is set down. For the step detection, the algorithm iterates through the window of local angular velocities of the y-axis for both leg-placed IMUs. If a pair is found that meets the conditions $\omega_{B,y,k} > \omega_w$ and $\omega_{B,y,k+1} \leq \omega_w$ a step is detected. The step threshold ω_w was set to $\omega_w = -60°/s$ empirically. In addition, the current counter of the respective NilsPod is saved so that only the new samples are evaluated during the next cycle.

Facing and Turning

The facing direction is calculated for both leg-based NilsPods individually by their reference orientation q_{ref} and their current orientation q_k. To only compare the heading value around the global z-axis W_z the two local vector $\overrightarrow{o}_{B,1}, \overrightarrow{o}_{B,2} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$ are transformed by the quaternion orientations respectively yielding $\overrightarrow{o}_{ref}$ and $\overrightarrow{o}_k$. Then the angle α_z between both around W_z is calculated.

$$\alpha_z = atan2(o_{ref,x}, o_{ref,y}) - atan2(o_{k,x}, o_{k,y}) \tag{3.12}$$

If the value of at least one leg-based IMU is $\alpha_z > 60°$ the facing direction is changed to *left*, if the value is $\alpha_z < -60°$ facing direction is *right*. In case both is true, it is *ambiguous*. Else it is *straight*.

The *turning* detection itself runs with the other movement detections every 0.5s. It is then checked if one leg-based NilsPod changed from a valid direction (*straight, left or right*) to the designated direction since the last update.

Reaching up

Reaching up is detected by a combination of three conditions for the data of the respective arm-based IMU:

1. The difference between the maximal vertical position $x_{z,max}$ at index i_{max} and the minimal vertical position $x_{z,min}$ at index i_{z_min} must be higher than the empirically set $x_{ru} = 0.7\ m$. The minimal vertical position is obtained after the maximal position in the range of samples before the maximal value.

2. To differentiate movement from e.g. going back up after a *picking up* movement, the average between zero-normed maximal position $\bar{x}_{z,i_{max}}$ and minimal position $\bar{x}_{z,i_{max}}$ needs to be positive. Zero-normed values are obtained by $\bar{x}_{z,i} = x_{z,i} - h_{diff,i}$ as the global vertical position $x_{z,i}$ is dependent on the difference between altitudes.

3. The direction of the hand is assessed. The angle between the global gravity vector $\overrightarrow{g}$ and a local vector $\overrightarrow{o}_{B,3} = \begin{bmatrix} -1 & 0 & 0 \end{bmatrix}^T$ transformed by the quaternion q_{max} needs to be $\alpha_{RU} > 120°$.

Picking up

To detect *picking up*, the same criteria as for *reaching up* are used in an adapted form. The difference between the minimum and maximum value is calculated. $x_{z,min} - x_{z,max}$ must be smaller than empirically set $x_{pu} = 0.45\ m$. Furthermore, the difference of the normalised positions must result in a negative value and the angle between the global gravity vector and the direction vector of the hand must be $\alpha_{PU} \leq 60°$.

Boxing

To detect *boxing* in a window three criteria need to be fulfilled for the data of the arm-based IMU:

1. The maximum of the gravity-removed local acceleration $a_{B,grm}$ in direction of the local x-axis needs to be greater than the empirically set $a_{bx} = 15\ m/s^2$. This refers to the rapid deceleration at the end of the boxing process. The comparatively small value is meant to account for the lower strength of the target group. To obtain the gravity-removed local acceleration, the gravity vector $\overrightarrow{g}$ is transformed by the quaternion inverse to the body frame for each sample and subtracted from the local acceleration value.

2. To limit the stroke direction to the XY-plane, the magnitude of the global acceleration for the same sample needs to be higher than a_{bx} as well.

3. The angle between the strike direction and global gravity $\overrightarrow{g}$ needs to be $60° < \alpha_{bx} \leq 120°$.

Clapping

For a *clapping* detection the following two criteria need to be fulfilled:

1. The maximum magnitude of the gravity-removed local acceleration $a_{B,grm}$ in the direction of the local B_{xz}-plane needs to be greater than the empirically set $a_{cp} = 30\ m/s^2$.

2. Criterion 1 must be fulfilled for both arm-based IMUs and the maximum number of samples between both maxima is $j_{cp} = 0.05/T_s$.

Jump move

The *jump move* is recognised as follows. The thresholds were set empirically.

1. First, there must be a maximum global velocity $\dot{x}_z > 1\ m/s$. This is the final jump motion (without leaving the ground).

2. In the period before that there must be a minimum negative global velocity $\dot{x}_z < 0.5 m/s$. This is caused by the knee bend.

3. At any point, there must be a maximal global angular velocity in W_x direction. This must be $\omega_x > 200°/s$. This is achieved by the arm swing at the end of the jump move.

4. Criteria 1-3 need to be fulfilled for both arm-based IMU.

Stepping over

Stepping over is detected recognized with the maximum global position $x_{z,max}$ of the leg-based IMU. As the empirical threshold with $x_{so} = 0.25\ m$ is comparatively low compared to the presumed high relative inaccuracy of the barometer, the maximum and minimum are not taken from the whole window, but since the last step started within the window. This is possible as the beginning of the *stepping over* movement is also detected as a step.

Stepping left or right

For the detection of sidewise stepping the order of leg movement is decisive. First, for both leg-based IMUs the maximum global velocity $\dot{x}_{x,max}$ and minimal global velocity $\dot{x}_{x,min}$ in W_x direction are analysed. For *stepping right*, both leg-based IMUs need a value $\dot{x}_{x,max}$ higher than the empirically set $\dot{x}_{sr} = 0.5$. In addition, the maximum on the right foot must be before the left foot for the foot sequence to be correct. For a *stepping left* to be detected $\dot{x}_{x,min}$ needs to be smaller than $\dot{x}_{sl} = -0.5$. The foot sequence here starts with the left foot.

3.1.4 Gameplay Development

The following part gives insights into the development of the gameplay and design choices for the game.

Cognitive training

In collaboration with gerontologist Daniel Schöne and psychologist Linda Becker, special emphasis was placed on the specific areas of cognitive training that can be implemented in the exergame during the development process. Regarding the cognitive domain to be trained, a distinction can be made between reaction time, processing speed, memory, verbal fluency, executive functions, attention and visuo-spatial abilities, each with multiple subcategories [Kue12]. For this exergame implementations for the following domains were considered:

- **Attention:** Attention is the process through which a person targets or concentrates on certain auditory or visual stimuli in the environment [Kue12].

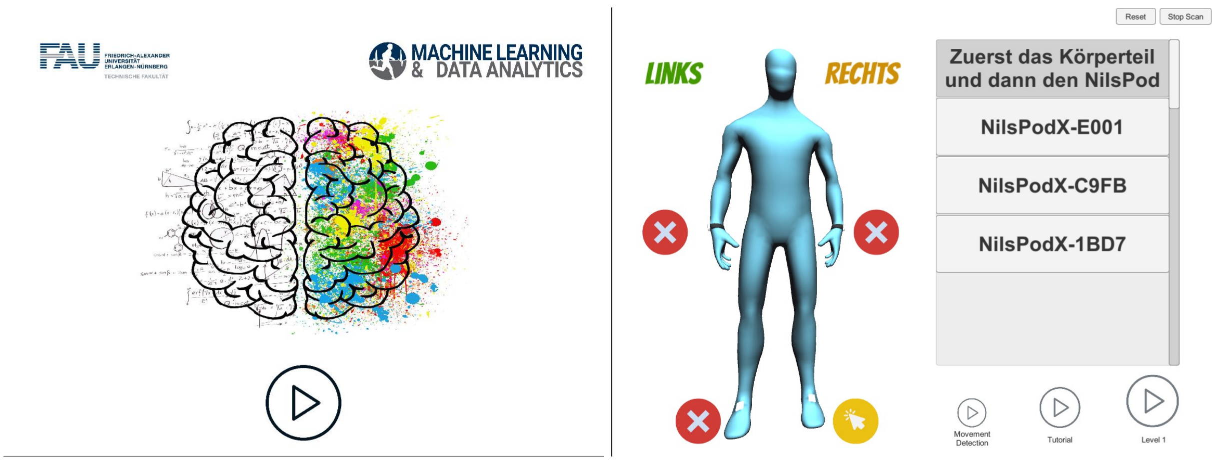
- **Visuo-spatial:** Visuospatial capacity is an aspect of visual perception that allows the orientation or position of objects in space to be processed mentally [Ira11].

- **Memory:** Memory is the capacity to retain, store and access the information. This ability is further divided into working memory, episodic memory and visual memory, among others [Kue12].
- **Executive functions:** Executive functions is a broad term that incorporates a number of different abilities. These include planning, cognitive flexibility and abstract thinking [Kue12].
- **Processing speed:** Processing speed describes the capacity to swiftly process information [Kue12].

The results regarding the implementations of the individual cognitive domains are described in section 4.1.

Game structure

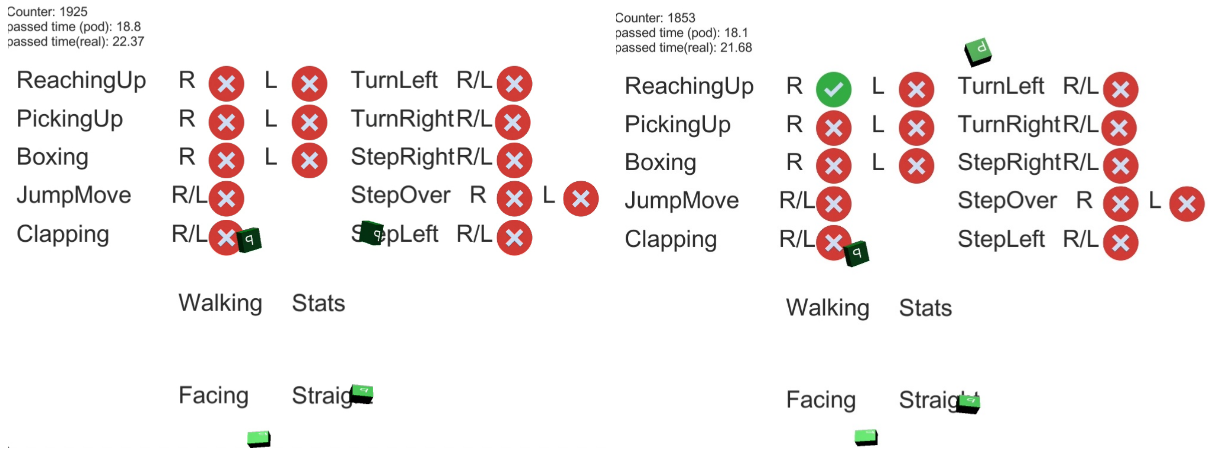
Four unity main scenes were implemented in the *MCI-Exergame* prototype: The *menu* (Figure 3.4), the *detection view* (Figure 3.5), the *tutorial* and the *level 1* (Figure 3.6). Furthermore scenes to visualize CSV-imported data, to start the evaluation and to visualize a single IMU were developed. The latter scenes are not part of the exportable application and can only be reached in the Unity development environment.



(a) Start screen (b) IMU selection screen

Figure 3.4: Displays the *menu*. Figure A shows the start screen when the app is started. Figure B visualizes the selection of pods for the according placement.

The game always starts with the menu. According to the current state of development, the menu is the only time that the player has to interact with the touchscreen of the device. During

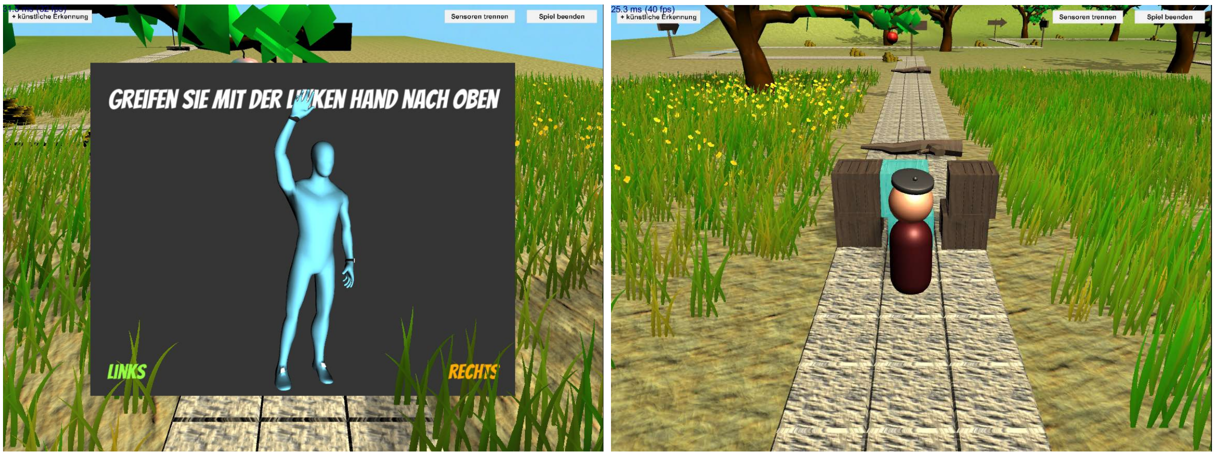


(a) *Detection view* scene before *reaching up* (b) *Detection view* scene during *reaching up*

Figure 3.5: Displays the *detection view* scene. The *detection view* is for development and visualization purposes. It shows all four pods with continuously updated orientation and position. Figure (a) shows a state before the *reaching up R* was started. Figure (b) visualizes the scene during the movement with a detection shown for *reaching up R*.

development, therefore, care was taken to ensure that the interface is very intuitive and responsive. Here, available NilsPods are searched for and can then be assigned to the corresponding body parts. At the moment of assignment, the corresponding calibration file is loaded from the memory. When all NilsPods are selected the user can select one of the three options in the menu to either navigate to the *movement detection* scene, the *tutorial* or the *level 1*. As soon as the corresponding button is pressed, the application changes to that scene and starts the connection process with the selected NilsPods. For this purpose the Unity Asset ”Bluetooth LE for iOS tvOS and Android” [Sha20] which implements the scanning, connecting to BLE devices as well as the commanding and subscribing to BLE services and characteristics for iOS, macOS or Android devices from one code base.

As the movement recording study was conducted before the main development of the *MCI-Exergame* It had already been determined that both usability and movement detection require a high level of instruction of the movements. Both the *tutorial* and the *level 1* consist of a parkour with the same movements. However, both scenes differ substantially in their objectives. In the tutorial, the player is supposed to get to know the game on his own. Due to high amount of non-valid or ambigious movements in the movement detection study, a higher specification of movements was needed. In order for the user to know which movement is required for which object, an instructor appears and shows the corresponding movement as a 3D animation. This disappears again as soon as the movement is detected or the user has moved too far away from



(a) Tutorial with instructor (b) Level 1

Figure 3.6: Displays two captures from the game. Figure A shows the explanation of the instructor for *reaching up L*. Figure B visualizes the scene during *level 1* where the player needs do a *boxing* move to destroy the boxes in front of the avatar

the object. In the tutorial, each object is interacted with once and therefore each movement is introduced. To progress in the game, the player must walk on the spot. This is also shown by the instructor in the *tutorial*. *Level 1* includes a significantly longer path which is designed as a circuit. In addition, the instructor only appears when a person remains standing for a longer than 5 seconds at an interactable object. *Level 1* Furthermore, the level does not end after a certain number of objects, but is completed 10 minutes after the start of the level.

Design and Interaction

The objective in terms of design and interaction in the game was to create an environment that was as pleasant and familiar as possible. With this in mind, an attempt was made to design a scenic outdoor world reminiscent of a walk and to make the objects for interaction as intuitive as possible. For each movement a specific interaction was designed and for all movements except *boxing* and *reaching Up* a distinction was made between right and left. *stepping left* and *right* was implemented and can be used to change lanes at any time in the game, but serves no further purpose in the gameplay. To move forward the player needs to walk on the spot. Whenever a game object is interactable it changes its color to a light blue or starts glowing light blue. After the correct movement is recognized the obstacle disappears. Depending on the object a sound is played as auditory feedback for the user (e.g. the clicking of coins after *picking up*).

All sounds, animations and 3D objects, except for the basic mesh of the instructor [CG_16]

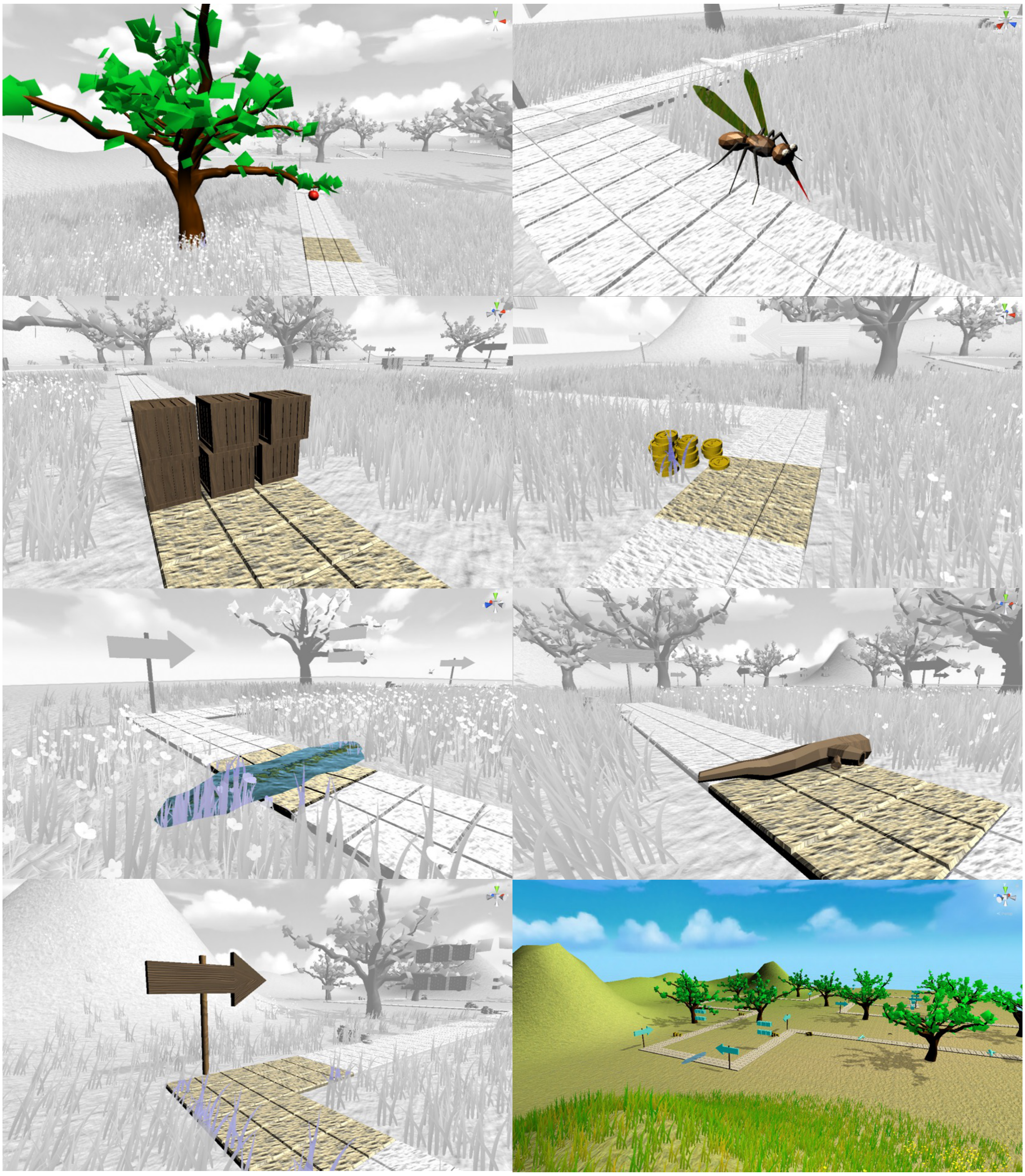


Figure 3.7: Main game objects in isolated mode and terrain view of *level 1*. In reading direction: apple tree for *reaching up*, mosquito for *clapping*, boxes for *boxing*, coins for *picking up*, river for *jump move*, tree log for *stepping over*, corner with arrow for *turning* and territory for *level 1*

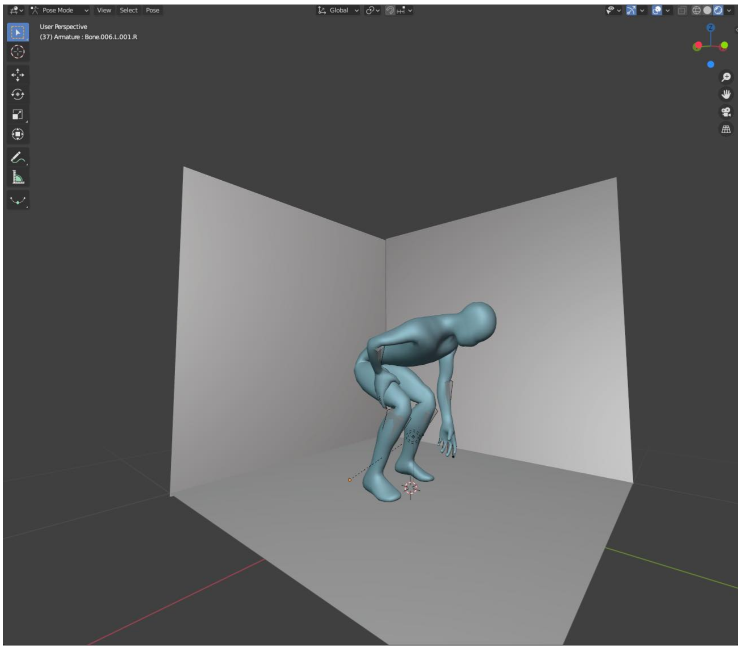


Figure 3.8: Blender environment showing the development of the *picking up L* movement for the instructor.

and the menu start [Eli17], were created by the author. The basic mesh for the instructor, shown in the menu, tutorial and level 1 was further developed in Blender to incorporate a skeletal model for the movement animations (Figure 3.8). Additionally, multiple materials and textures are used from the sources stated in the beginning of this chapter.

3.1.5 Output Measures

For evaluation of the game and player performance three logs are provided by the MCI Exergame and are saved to the local storage of the game device in CSV format. The saved data is easily accessible in the data folders of iOS or Android and can be retrieved by cable or wireless technology. The logs are saved when the *end game* button is pressed and as a recovery file every two minutes during playtime. The *GameLog* includes all major events in the game separated into categories: *Main*, *Detection*, *Gameplay* and *BLE*. Every entry is saved with a real timestamp. Step detections are also included here. The *DetectionLog* includes every run of the detection system and a timestamp since the game start for every log. Additionally, for a detection the real counters of the decisive pod are provided as well as the information if it was a *first detection*, meaning that it was the first window that detected this movement. Only these *first detections* count in the gameplay. Additionally, there is an IMU data log file for every pod. These provide all raw data from the point, where all pods were connected as well as the normalised counter. The data structure in the *IMULogs* is the same as the CSV export of recordings in the Portabiles App [por21].

3.2 Movement Recording Study

3.2.1 Purpose of the Study

The primary purpose of this first study was to create a basis for the development of movement recognition for elderly people. This should provide insight into the associated movements and their execution as well as an understanding how they should be instructed. The aim was to record as much diversity of associative movements as possible. From this pool of movements, the sequences of movements that were to be recognised should then be identified and a solution should be found to visualise them clearly in the game. The developed visualisation and recognition should then be tested in a second study in which the participants were given more specific instructions. Due to the COVID-19 pandemic, a second movement recording on elderly people with corresponding instructions from the game could not take place. Therefore, the developed movement recognition was also tested on selected data from this study. The movement recording study was scheduled at the beginning of the development of the MCI exergame.

3.2.2 Participants

The participants were all healthy (for their age) community-dwelling older people with a sufficient level of English. The latter was due to a restriction of the ethics approval. Five participants were found which matched the given criteria. The participants' average age was 73.0±10.1 years with a minimun of 60 and a maximum of 83 years.

3.2.3 Procedure

In order to obtain a diverse set of movement recordings for the development, the study was divided into three phases. In each of the phases, the participant had a different association with the movements. In addition to the three phases, the participant was required to walk on the spot for 30 seconds at the beginning of the study.

In phase I, the participant was given simple instructions for a movement, for example: ”Reach up with your right arm”. In phase II, this instruction was linked to an instance of the real-world. For example, ”Reach up with your right arm to pick an apple hanging above your head”. In phases I and II, the same movement was performed five times in succession, always on the study instructor's start signal. The time interval between movements of the same kind was about five seconds.

The order of phases I and II as well as the movements within the phases were randomized. In phase III, the participant performed all movements in random order. Every movement was performed at least once. In contrast to phases I and II, the participant was supported by a visual picture for the movement that was displayed with a projector on the wall B. The starting signal for a movement was the change of visual input. Further details on the study procedure can be found in Appendix B.

To record the movements of the participant, four IMUs were placed as described in Section 3.1.2. All movements were recorded with the *Portabiles* app running on an Android-based smartphone. Additionally, footage of the participants movements was recorded.

The study took place under the shared supervision of Daniel Schöne and the author of this thesis.

3.2.4 Data Analysis

First, the videos were manually annotated. The exact moment of the following events was entered: Start of the movement recording in the Portabiles app, start signals of the study leader. By converting the times as the difference to the start of recording of the IMUs and multiplying them by the sampling rate $f_s = 102.4\ Hz$, the movement data could be assigned accordingly. In the case that the start time of the movement recording was not visible in the Portabiles app, this was estimate by using later movements of the partipant and calculating these back to the start of the recording.

For the following analyses, data were excluded for the following reasons:

1. **Missing IMU data:** One or more of the data files were missing, corrupted or overwritten by the Portabiles app.

2. **Missing video:** Video footage of the movement was not recorded.

Due to the COVID-19 pandemic, it was not possible to test movement detection separately with the visual guidance from the game with elderly people. Nevertheless, in order to evaluate the motion detection in an initial step, it was tested on the recordings of this study. For this purpose, a scene was created in the MCI Exergame, which can automatically read IMU data in CSV format and outputs at what time interval which detection was made. The interval consisted of the detection time interval described trimmed window described in the beginning of Section 3.5. These detections were then compared with the start signals from the video annotation. As the game runs its movement detection every 0.5 seconds over last 3 seconds, most movements are

movement	minimal criteria
reaching up	specified hand reaches up over the head
picking up	specified hand reaches down below knee
boxing	visible acceleration of specified arm in frontal direction
clapping	fast merging of both hands
jump move	moving the upper body down and up again by bending the knees, while going upwards, swinging the arms upwards
stepping over	a higher step than is needed for normal walking
turning	rotation of at least one part of the lower body
stepping left/right	a step with each foot in the specified direction

Table 3.2: Displays minimal criteria for the movements of the movement recording study to be classified as valid.

detected more than once. While in the game, the first detection is used to support a continuous game flow and other detections of the same movement are ignored, in the evaluation the start of the last detection of the movement is evaluated. This had the advantage that the movement starts could be better assigned to the video annotations. As it was the last time the detection was made, the start of the detection interval was most closely coincided with the real movement start. Movements were counted as detected if the movement start t_{MS} was in the correct interval compared to the start signal by the study leader t_S. The interval was set to $t_S - 1 < t_{SM} < t_S + 4$.

As the system as well as the evaluation yield a multi-label binary classification per instance, individual confusion matrices for each movement were calculated with the one-vs-all approach. Additionally, *precision*, *recall* and *F1-score* as harmonic mean for each movement type were calculated. Other calculated metrics for overall movement analysis include: *Hemming loss*, *Exact Match Ratio* as well as Micro and Macro averaging over the metrics mentioned above [Moh10, Tso07].

Exact Match Ratio provides the strictest measure in the multi-label evaluation. Here a prediction vector of a sample is only true, when all predictions are correct. The Exact Match Ratio with the ground truth Y_i and the prediction vector Z_i as calculated was follows:

$$EMR = \frac{1}{n} \sum_{i=1}^{n} I\left(Y_i = Z_i\right) \tag{3.13}$$

In contrast, the Hemming Loss also takes predictions into account that are partially correct. Here, the share of all labels of all predictions that was incorrect is determined. The lower the Hemming Loss, the closer the predictions are to the ground truth. Here, N is the number of evaluated instances and L the number of labels.

$$HL = \frac{1}{N \cdot L} \sum_{i=1}^{N} \sum_{j=1}^{L} \operatorname{xor}\left(y_{i,j}, z_{i,j}\right) \tag{3.14}$$

While macro-averaging describes the averaging of values such as precision and recall across the different classes, micro-averaging is the averaging over the individual instances. Depending on the distribution, classes with more instances can have a greater influence in micro-averaging, while classes with fewer instances can have a disproportionately large influence on the scores in macro-averaging. [Moh10, Tso07].

3.3 Usability Study

3.3.1 Purpose of the Study

The purpose of the following study was to test multiple aspects of the exergame's usability. The usability is essential for independent use of the game at home. In this second study, the focus laid on the factors of the system's usability, gaming experience and the enjoyment of physical activity especially. These factors were tested on the final state of the developed *MCI-Exergame* prototype. Additionally, physical measures of the participants as well as game statistics were collected during the interaction with the exergame. Supplementary information including a step-by-step implementation can be found in the study protocol of the study in Appendix C.

3.3.2 Participants

The initial intention was to test these elements of the developed MCI Exergame with healthy older people. This would have the advantage that both the enjoyment and interaction with the game as well as the physical and cognitive load would have been tested directly on the target group. Due to the current COVID-19 pandemic, in which older people in particular were affected by severe courses, the study could not be tested on an elderly population. In accordance with the supervisors and advisors of this Master's thesis, it was then decided that the same study procedure should first be tested with younger subjects instead. As the study was designed to be suitable for elderlies in Germany, the study was conducted in German to reduce language barriers.

Seventeen healthy younger participants conducted the study. Their subject characteristics are displayed in table 3.3.

Number	17
Age[years] ± SD	24.6 ± 3.0
Height [cm] ± SD	174.3 ± 10.8
Weight [kg] ± SD	72.2 ± 11.8
Sex (m/f)	6 / 11
strong hand (r/l)	16 / 1

Table 3.3: The subject characteristics of the usability study.

3.3.3 Procedure

The procedure of this study was planned with the intent of a broad evaluation of the exergame's usability. Due to the COVID-19 pandemic, the usability could not be evaluated as planned. Instead of testing the exergame on elderly people, only a test on young people was possible. In addition, specific measures were taken to reduce the risk of transmission to a bare minimum during the study. This included, among other things: Large distances between study supervisor and participants, an FFP2 mask requirement, thorough disinfection after each usage of the sensors, and contact tracing. An iPad 2017 with iOS 14.4 functioned as the mobile device for playing the MCI-Exergame. With a projector, the game was shown to the participant on the room wall. 4 NilsPods from Portabiles GmbH [por21] were used and attached to the corresponding body parts as described in Section 3.1.2. The participants were filmed for later video analysis and the HRV Logger App [Mar] was used to record heart rate metrics. Figure 3.9 shows the main steps of the study. These will now be described in brief. For a more detailed description of the individual steps, see the study protocol in Appendix C.

First, the participant was informed about the study and the background of the study. The participant filled in the documents *Profile of Subject*, *Physical Activity Readiness Questionnaire* and *Declaration of Consent*. The Polar HRV sensor and the NilsPods were then placed on the designated body areas and the subject received instructions for the first phase of the test. These included, among others, an explanation of the think-aloud method and interaction with the game in the *tutorial* phase. With the latter, it is worth noting that the subjects were told that they should wait three seconds before the next trial if a movement was not detected properly and that if a correctly executed movement was not detected three times in a row, an artificial detection is generated by the study supervisor on the tablet directly. This was implemented in order not to unnecessarily hinder the flow of the game with problems in recognising individual movements. The app was then started on the tablet by the study supervisor, the IMUs were chosen depending on their location and the *tutorial* was started. The instructions for the next phase, *level 1*, were the same except for the explanation of the Borg Scale of Perceived Exertion. While the participants

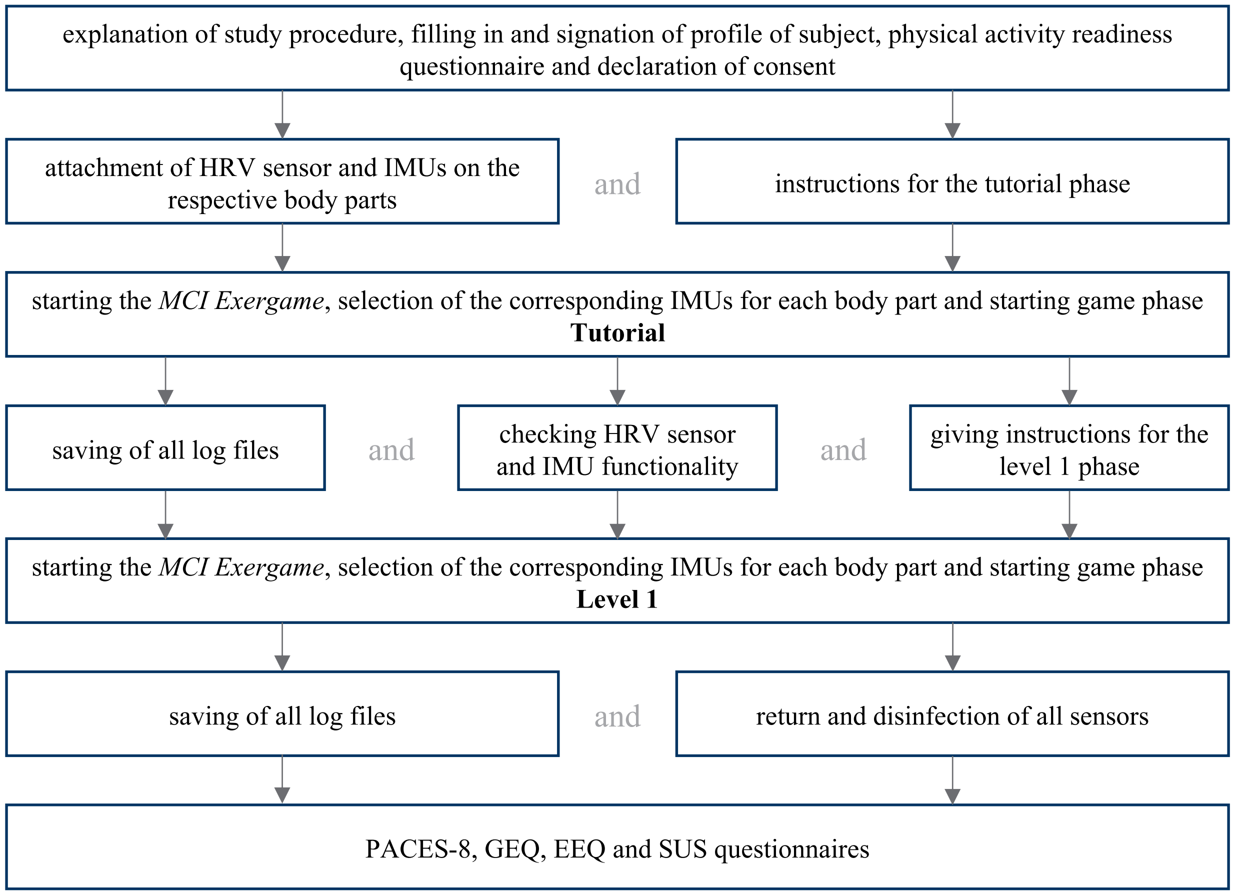


Figure 3.9: An overview of the main steps of the usability study flow.

played *level 1* they answered the Borg Scale 3 times. After completing both stages the sensors were returned and disinfected and the participant was asked the open questions and filled in the 8-item physical activity enjoyment scale (PACES-8), game experience questionnaire (GEQ), exergame experience questionnaire (EEQ) and system usability scale (SUS) (Appendix C).

3.3.4 Measurements

During and after the two-staged experiment, different quantitative and qualitative information was collected. The quantitative measurements included recorded movement data, logs on the operation and usage of the app, heart rate and heart rate variability as well as the Borg rating of perceived exertion scale. Both, heart rate variability (HRV) and Borg scale were used to examine the participants' physical response during the exergame. Additionally, participants were asked to fill in the SUS, the PACES-8, the GEQ and the EEQ to evaluate their enjoyment during the exergame as well as their perceived usability of the system. In addition, qualitative data were recorded. These include the think-aloud statements of the participants and the answers to the open questions. In the following, the details of the recorded information are provided for each measure, respectively.

Game Logs

The evaluation of the game logs (Section 3.1.5) can be helpful in assessing the interaction with the game and the associated usability. The evaluation also provides information about the technical implementation of the game and the movement detection. In addition, the evaluation of the IMU signals can help to draw conclusions about the information on physical activity. For the purpose of this study, the detection system of the game ran for all movements at all times and not only when needed for an interaction with a current obstacle. The evaluation of the GameLogs allows the calculation of these results:

1. **Amount of Movement**: The IMU raw values were used to calculate the *Amount of Movement* of each individual pod. For this purpose, the calculations used by the R-based open-source project GGIR [Mig19, Row16] were carried out with minor deviations in Python. Instead of the autocalibration[van14] performed in the GGIR project, the existing Ferraris Calibration by Portabiles for each NilsPod was used on the raw data [Fer, por21]. The calibration increases the accuracy of the measured values by adjusting them according to the individual IMU (Section 3.1.2). Afterwards, the average magnitude of dynamic acceleration $\bar{a}$ was calculated, by using the Euclidean Norm on each acceleration vector and subtracting $g = 9.81$.

$$\bar{a}_k = \sqrt{x_{k,x}^2 + x_{k,y}^2 + x_{k,z}^2} - g \tag{3.15}$$

   Afterwards, negative values were rounded to zero and 5s epochs were aggregated

$$\bar{a}_k = \begin{cases} \bar{a}_k & if\ \bar{a}_k \geq 0 \\ 0 & else \end{cases} \tag{3.16}$$

$$\bar{a}_{5s,i} = \sum_{k=5i}^{5i+4} \bar{a}_k \tag{3.17}$$

2. **Interaction Metrics**: From the interaction starts and ends in the *GameLog*, interaction times were calculated. The interaction time was defined by the time an object became

interactive until the interaction ended. This was either due to a detected interaction of the user with the object or the passing-through of the player without interacting. Data entries with interaction times of more than 20 seconds and less than 0.1 seconds were excluded, as these are most likely caused by an incorrect or missing entry in the *GameLog*. Additionally, the percentage of movements that were detected in the first four seconds since interaction start was evaluated. From the *GameLog* and the video footage, the number of movements that needed to be repeated as well as artificial detections created by the study supervisor were obtained. The repetition rate movements only counted when they were correct and started after the interaction start signal in the game. No exclusion was necessary here, as the analysis of the video was not affected by the above-mentioned *Gamelog* problem.

3. **Step Count** The number of detected steps of each participant were retrieved from the *GameLog* as well.

Heart Rate Metrics

The measurement of the heart metrics aims to classify the level of physical activity of the participants during the study. Heart rate metrics were recorded with a Polaris heartrate H7 sensor and the App *Heart Rate Variability Logger*[Mar] during both stages of the gameplay. For the analysis, the HR evaluation software *KUBIOS* [Kub] was used to calculate overall and minute-wise heart rate (HR) and root mean square of successive RR interval differences (RMSSD) for each participant in Level 1.

In-Game Data Collection

During the *tutorial* and *level 1*, the participant was instructed that any thought regarding the current situation should be spoken out loud. Using the think-aloud method [van94], qualitative situational data can be collected, which, unlike questions that are answered afterwards, can represent a progression. Additionally, while playing *level 1* the participant was asked 3 times about his level of perceived exertion. For this purpose, an adapted Borg Rating of Perceived Exertion Scale [BOR82] was used. This adaptation (Appendix C) is an interval-scale from 0-10 with a German translation provided by the Institute of Medical Physics at Friedrich-Alexander University Erlangen-Nuremberg.

Questionnaires and Open Questions

The particular questionnaires were chosen for the resulting broad spectrum of usability aspects in an exergame. This should reliably test the factors physical enjoyment, game enjoyment, and usability of the exergame. To obtain more direct measurements and answers, additional open questions were implemented (Appendix C).

The PACES-8 is based on the original 18-item Physical Activity Enjoyment Scale[Ken91]. The shortened 8-item Physical Activity Enjoyment Scale [Mul11] was later developed and validated on older adults. As with the original PACES, the objective was to give researchers and clinicians a tool to determine the enjoyment of physical exercise quickly and accurately. The German version of the PACES-8 was provided by the Institute of Medical Physics at Friedrich-Alexander University Erlangen-Nuremberg (see Appendix C). To complete the PACES-8 the participant had to fill in eight items on a bipolar survey scale with seven intervals. In the evaluation, each item was given a score between 1 to 7 (7 to 1 if the item was reverse-coded) and summed for each participant.

The SUS[Bro96] is a simple and quick way to measure a device's usability from the user's perspective [Bro13]. On a scale of 0-100, the SUS score indicates the usability performance in terms of effectiveness, efficiency, and overall ease of use. Scores below 68 are rated below-average, whereas results over 68 are considered above-average. Due to its widespread use, the SUS offers a good opportunity to assess the usability of one's own product and to compare it with others. The SUS consists of 10 items which can be rated on a scale of 1 - strongly disagree to 5 - strongly agree. The German version of the SUS, which was used in this study is a translation by the Usability Experience department of SAP[Rei] (see Appendix C) . For the determination of the system usability score, the methods described by Brooke were used. (see [Bro96]). Addionally the Pearson's correlation was calculated between the SUS score and both interaction metrics individually.

Another questionnaire that was completed by the participants, was the GEQ[W.A13, K. 07]. The Game Experience Questionnaire is a self-report instrument designed to assess the experience of playing digital games comprehensively and accurately. The GEQ is separated into three independent modules: 1. The core questionnaire, 2. The post-game questionnaire and 3. The social presence module. In the scope of this study, only the core questionnaire was utilized as the exergame did not fit criteria for the social presence module and the post-game questionnaire is recommended if the playing decisions happen naturally (e.g. when to start or stop). The questionnaire evaluates the gaming experience in seven categories: *Immersion*, *Flow*, *Competence*, *Positive and Negative Affect*, *Tension* and *Challenge*. The participant states his or her acceptance of 33 statements on a scale from 0 - not at all to 4 - extremely. The German translation of the questionnaire was provided by the chair of Health Psychology at Friedrich-Alexander University Erlangen-Nuremberg (see Appendix C).

The EEQ[Fit20] is a new measure specifically designed for the evaluation of enjoyment in exergames. Therefore it incorporates multiple questions from game experience questionnaire, Physical Activity Enjoyment Scale and the Immersion Experience Questionnaire[Jen08]. The user needs to fill in 20 questions on a 5-point Likert scale from strongly disagree to strongly agree. As EEQ is only validated in the original study and used in very few studies at the time of this study, it is used as an addition to more comparable measurements like GEQ and SUS. The German translation of the EEQ provided by the chair of Health Psychology at Friedrich-Alexander University Erlangen-Nuremberg as well (see Appendix C) Due to the falsely reverse-coded translation of Question 4 of the EEQ the scores for this element were again reverse-coded before the evaluation. Besides this, the EEQ score was calculated as described by Fitzgerald et al. [Fit20].

Chapter 4

Results

In chapter 4, the results of the exergame development and the conducted studies are presented. Separate sections deal with the development of the exergame, the conducted movement recording study, and the usability study.

4.1 MCI-Exergame

In this chapter, the resulting *MCI Exergame* prototype and its properties are described. Limitations of the current state will be discussed in Section 5.1.

The final state of the *MCI Exergame* is a Unity project which supports the platforms Android, iOS, and macOS. The prototype connects reliably with 4 NilsPods one after the other. Although the development project in Unity allows the import of CSV files with a sampling rate of 102.4 Hz to visualise or evaluate movements, only NilsPods with a set sampling rate of 51.2 Hz can be connected for live operation. The first data analysis and calculation of parameters were then performed on each incoming sample directly, yielding parameters including acceleration, velocity, position, and orientation. All relevant parameters are then stored into the buffer with a length of 3 seconds. Every 0.5 seconds the movement detection is performed on basis of the current buffer and every 0.2 seconds step detection evaluates unseen samples. The playable exergame is designed as a parkour where different movements are needed to interact with obstacles in the game. The following features were implemented for the different cognitive domains.

For this exergame implementations for the following cognitive domains were considered:

- **Attention:** New objects appear (boxes, mosquito), which needs to be reacted to. Additionally, short-term reaction is needed an object becomes interactive.
- **Visuo-spatial:** Visuo-spatial abilities are required by the exergame itself, especially with the 3D surrounding. While many objects are visible for a longer time, they can only be interacted with when you are close enough in the virtual 3D world.
- **Memory:** No direct memory task was implemented.
- **Executive functions:** Inhibition was implemented in a sense that objects (even appearing objects) can only interacted with as soon as game allows it.
- **Processing speed:** No direct processing speed task was implemented. Processing speed is inherently trained by exergame and complex 3D environment itself.

The game has been tested on an iPad 2017, an iPhone 6s, a MacBook Pro (early 2015), and a Google Pixel 4, each with the device's native resolution. While the app could be successfully started and the game could be played on all devices, the MacBook Pro had the most performance problems, which became apparent in the low framerate.

4.2 Movement Recording Study

The movement recording study aimed to generate a data basis for the development of the *MCI-Exergame* prototype. In addition, an initial evaluation of the movement recognition was carried out on the basis of valid movement data. The results of this study therefore make a distinctive contribution to the assessment of the feasibility of an exergame for mild cognitive impairment (MCI) and can provide initial answers to the first research question concerning the detection of movements.

4.2.1 Movement Validity

In the course of the study, 150 inertial measurement unit (IMU) recordings of movements were obtained from five elderly people. These were composed as follows. Five recordings include sequences of 30 seconds of walking on the spot. Phases I and II each include 70 recordings. Each containing about five identical movements, resulting in 701 single movements. Five recordings belong to phase III. The phase III recordings contain a total of 148 individual movements.

Due to ”missing IMU data” and ”missing video” (Section 3.2.4), datasets had to be excluded from further analysis. From the total of 849 movements, 35 (4.12%) were excluded due to ”missing IMU Data” and 29 (3.42%) were excluded due to ”missing video”. Regarding further IMU problems, such as short-term strong peaks in the barometer, no data was excluded. It was argued that these can also occur in any situation and should therefore be taken into account.

The next step was to check to what extent the participant’s movement corresponded to the intended movement. This step was conducted by a video analysis. All movements which fulfilled the minimal criteria were considered valid (Section 3.2.4). A total number of 164 (20.84% of the non-excluded data) of movements were considered not valid for the intended movements (Table 4.1). An especially high percentage of invalid movements over all phases can be seen for *picking up R* (53.23%) and L (45.76%), for the *jump move* (49.09%) as well as *turning right* (45.00%) and *turning left* (48.28%). The percentages of invalid votes differ only slightly between the phases: Phase I - 20.11%, phase II - 18.97%, phase III - 18.24%.

4.2.2 Movement Detection Evaluation

As described in Section 3.2, the obtained recordings in this study were also used for an initial evaluation of the developed motion detection system. For this purpose, only movements that had previously been classified as valid were used. Additionally, due to the low number of valid

movement	non-excluded movements	invalid movements	invalid phase I	invalid phase II	invalid phase III
reaching up R	59	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)
reaching up L	40	4 (10.00%)	0 (0.00%)	0 (0.00%)	4 (40.00%)
picking up R	62	33 (53.23%)	16 (59.26%)	16 (61.54%)	1 (11.11%)
picking up L	59	27 (45.76%)	16 (61.54%)	10 (50.50%)	1 (7.69%)
boxing R	56	1 (1.79%)	0 (0.00%)	0 (0.00%)	1 (10.00%)
boxing L	59	1 (1.69%)	0 (0.00%)	0 (0.00%)	1 (7.69%)
clapping	56	5 (8.93%)	0 (0.00%)	5 (20.00%)	0 (0.00%)
jump move	55	27 (49.09%)	20 (100.00%)	5 (20.83%)	2 (18.18%)
stepping over R	43	4 (9.30%)	0 (0.00%)	0 (0.00%)	4 (57.14%)
stepping over L	55	1 (1.82%)	0 (0.00%)	0 (0.00%)	1 (12.50%)
turning right	60	27 (45.00%)	10 (40.00%)	15 (60.00%)	2 (20.00%)
turning left	58	28 (48.28%)	9 (39.13%)	15 (60.00%)	4 (40.00%)
stepping right	59	3 (5.08%)	0 (0.00%)	0 (0.00%)	3 (33.3%)
stepping left	66	3 (4.55%)	0 (0.00%)	0 (0.00%)	3 (18.75%)

Table 4.1: Shows an overview of the movements which are not valid and their distribution over the phases. Percentages are based on the non-excluded movements of the corresponding phase(s).

movements in some categories, it was decided to combine all three phases together for the following analysis.

The Hemming loss over all movements was 3.40% and the Exact Match Ratio was 63,69%. In the recall, which is particularly important for the flow of the game, especially high values were achieved for *reaching up* (93.22%, 94.44%), *clapping* (92.16%), *turning* (96.67%, 96.67%) as well as stepping *right* and *left* (96.43%, 90.48%). In contrast, *picking up* (65.52%, 65.62%) and *stepping over* (47.73%, 40.74%) achieved particularly low scores in this regard (Table 4.2).

The precision on the other hand shows considerably lower values. Especially *stepping over* (60.00%, 48.89%), *reaching up* (68.75%, 66.67%), *boxing* (66.67%, 74.60%) and *picking up* (70.37%, 75.00%) stand out because of their low precision. For these movements, in a second step, it was analysed how the false positive values distribute for the corresponding true movements (Table 4.3). Noticeably, *reaching up* was detected particularly often with *jump move* and *clapping*. *Picking up* was most frequently detected with *jump move* as the true value, but was more widely distributed. *Boxing* was most frequently detected when *reaching up* on the corresponding side. *Stepping over* was generally detected with various leg-based movements, but *stepping over L* in particular was most often detected with *picking up R* was the true value. In many of these cases, the false positive was detected together with the correct movement.

movement	precision	recall	f1-score	support
reaching up R	0.6875	0.9322	0.7914	59
reaching up L	0.6667	0.9444	0.7816	36
picking up R	0.7037	0.6552	0.6786	29
picking up L	0.7500	0.6562	0.7000	32
boxing R	0.6667	0.7636	0.7119	55
boxing L	0.7460	0.8103	0.7769	58
clapping	0.8704	0.9216	0.8952	51
jump move	1.0000	0.7500	0.8571	28
stepping over R	0.6000	0.4773	0.5316	44
stepping over L	0.4889	0.4074	0.4444	54
turning right	0.8889	0.9697	0.9275	33
turning left	1.0000	0.9667	0.9831	30
stepping right	0.7500	0.9643	0.8437	56
stepping left	0.8261	0.9048	0.8636	63
micro avg	0.7444	0.7978	0.7702	628
macro avg	0.7603	0.7946	0.7705	628

Table 4.2: Shows detection statistics for each individual movement as well as summarised as micro and macro average.

false positives	RU R (T)	RU L (T)	PU R (T)	PU L (T)	BX R (T)	BX L (T)	CP (T)	JM (T)	SO R (T)	SO L (T)	TR (T)	TL (T)	SR (T)	SL (T)	total count (TP+FP)
RU R (P)	0	0	0	1	3	0	9	12	0	0	0	0	0	0	**80**
RU L (P)	0	0	1	2	0	0	7	6	1	0	0	0	0	0	**51**
PU R (P)	2	0	0	1	1	0	0	2	0	0	1	0	1	0	**27**
PU L (P)	0	0	0	0	1	1	0	4	0	1	0	0	0	0	**28**
BX R (P)	14	0	0	0	0	1	4	1	0	0	0	0	0	1	**63**
BX L (P)	0	11	0	1	0	0	2	2	0	0	0	0	0	0	**63**
SO R (P)	1	0	0	1	0	1	0	2	0	2	3	1	3	0	**35**
SO L (P)	0	1	6	3	0	1	0	2	2	0	3	0	2	3	**45**

Table 4.3: Presents the false positive distribution for *reaching up*, *picking up*, *boxing* and *stepping over* movements. (T): True value, (P): Predicted value, RU: *Reaching up*, PU: *Picking up*, BX: *Boxing*, CP: *Clapping*, JM: *Jump move*, SO: *Stepping over*, TR: *Turning right*, TL: *Turning left*, SR: *Stepping right*, SL: *Stepping left*, TP: True positive, FP: False positive.

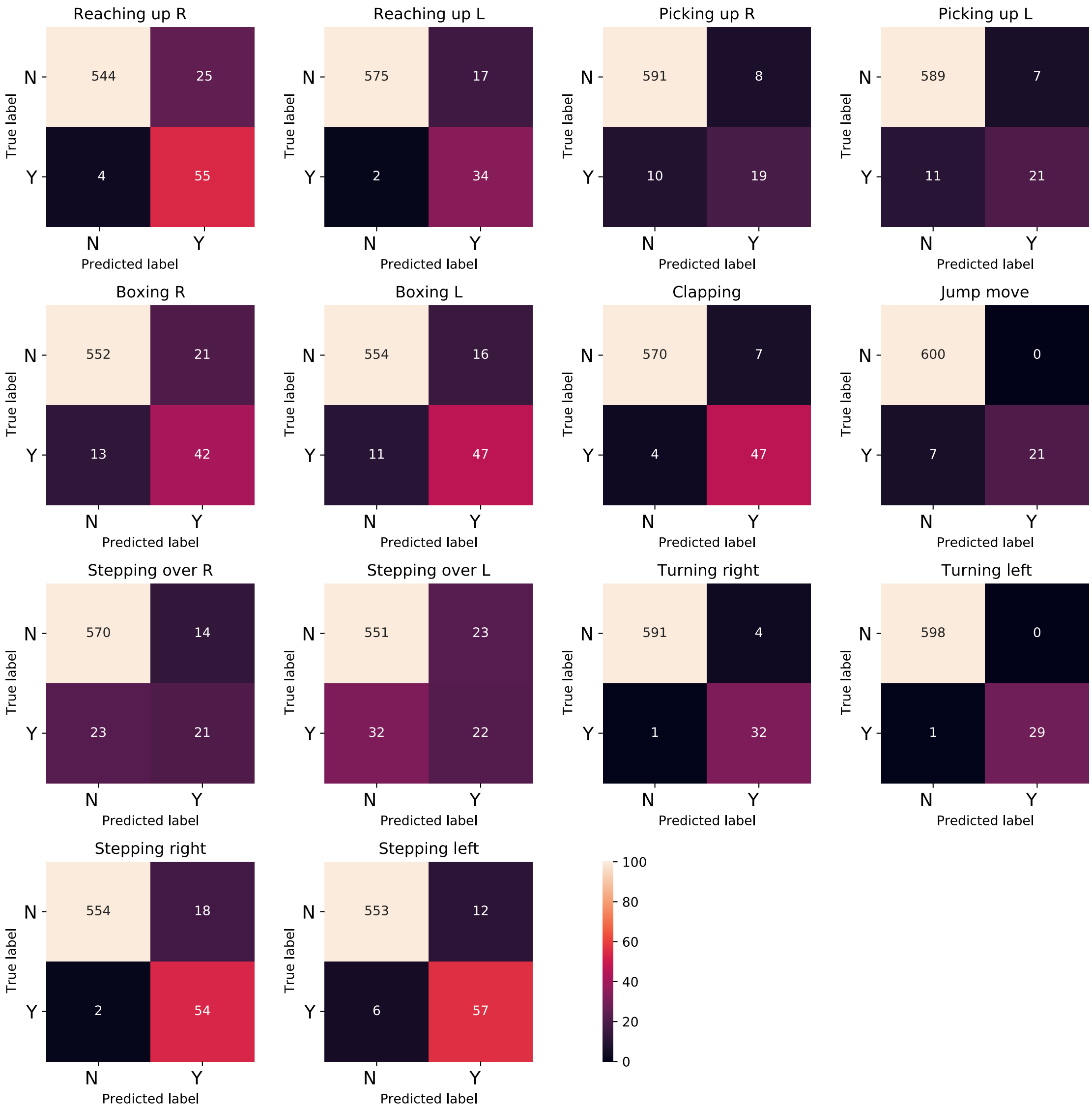


Figure 4.1: Shows one-vs-all confusion matrices for each movement. True negative: (0,0), false positive: (0,1), false negative: (1,0), true positive: (1,1).

4.3 Usability study

This section reports the results of the usability study. The results of the study are primarily concerned with the evaluation of different aspects of usability. In addition, interaction metrics indirectly provide information about movement detection in the game flow. Furthermore, physical metrics such as the heart rate can provide a first assessment of the person's exertion.

4.3.1 Physical Measures and Exertion

The following part presents the results on physical measures. First, the heart-related measurements are discussed and then the calculations regarding the amount of movement are presented. Tables 4.4 and 4.5 display minute-wise aggregations of HR and RMSSD as means over all participants, respectively. It should be noted that the standard deviation in both tables is based on the differences between minute-wise aggregations of each participant and is not based on all data points. In Appendix D and Appendix D these metrics are visualized for each participant. The mean heart rate (HR) participants 10 minutes measured was 95.45±16.24 beats per minute. The highest mean change in heart rate is from HR 1 to HR 2 by 4.08 beats per minute. Afterwards, the heart rate stays generally constant.

For an average age of 24.6 years, an intensity of 48.83% of the maximum heart rate is calculated. The intensity is thereby classified below moderate exercise (50% - 70% of maximum heart rate) [May21, Ame21]. Individually, 9 participants had a mean heart rate above 50% of their age-dependent maximum heart rate, classifying the exercise as moderate. All other participants' means were below 50% of their maximum heart rate.

HR	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**1-10**
mean	91.27	95.35	95.90	96.59	95.03	96.62	95.27	95.70	95.56	96.94	95.42
std	15.57	14.63	15.03	15.83	16.48	17.18	17.44	17.97	17.29	17.26	16.24
min	65.41	70.74	71.47	70.64	65.39	65.65	64.92	64.02	69.59	68.90	67.65
max	117.45	119.62	124.79	130.52	125.59	127.70	129.86	125.92	129.43	128.68	125.96

Table 4.4: The course of HR metrics over 10 min. Each column represents values of all participants for that minute, except for the last column which represents the mean over 10 min. All values are presented in beats per minute.

The mean root mean square of successive RR interval differences (RMSSD) value over all participants was 25.76±15.70 ms. The high standard deviation is also reflected in the individual data, whose values differ substantially. Compared to the mean calculated by Nunan et al. from 5-minute measurements of 42±15 ms, the mean value measured here is considerably lower [Nun10].

RMSSD	1	2	3	4	5	6	7	8	9	10	1-10
mean	24.74	23.44	25.64	26.34	26.48	22.58	22.44	27.74	25.44	24.93	25.76
std	18.11	16.29	15.26	16.19	15.59	18.24	15.67	19.35	16.96	16.62	15.70
min	4.99	6.79	7.64	7.66	6.39	5.28	5.17	6.84	5.18	5.18	8.81
max	71.83	75.91	69.75	68.12	69.65	85.09	70.32	83.04	69.91	71.85	74.02

Table 4.5: The course of RMSSD metrics over 10 min. Each column represents values of all participants for that minute, except for the last column which represents the mean over 10 min. All values are given in milliseconds.

The calculated amount of movement over the aggregated 5s-samples as described in Section 3.3 yields the following results over participants. The right leg IMU was moved $0.88 \pm 0.37\ m/s^2$, the left leg $0.82 \pm 0.27\ m/s^2$, the right arm $1.39 \pm 0.61\ m/s^2$ and the left arm $1.22 \pm 0.55\ m/s^2$ in average over participant means. This results in a lower extremity sum of $1.69 \pm 0.63\ m/s^2$ and an upper extremity sum of $2.61 \pm 1.14\ m/s^2$ (Figure 4.2). For these values, the standard deviations are based on the distribution of mean values in participants, not all calculated 5s-samples.

4.3.2 Interaction Metrics

movement	**obstacle count**	**obstacle count lower 4s**	**mean interaction [s]**	**SD interaction [s]**
reaching up R	98	96 (97.95%)	1.692	0.935
reaching up L	155	144 (92.90%)	1.881	1.507
picking up R	97	75 (77.32%)	3.538	3.609
picking up L	146	113 (91.10%)	3.241	2.955
boxing	191	180 (94.24%)	1.803	1.707
clapping	107	106 (99.07%)	1.667	0.663
jump move	126	111 (88.10%)	2.492	2.314
stepping over	132	108 (81.81%)	2.501	2.343
turning right	192	189 (98.43%)	1.794	1.355
turning left	112	110 (98.21%)	1.845	1.301

Table 4.6: The table represents the interaction time per movement type across all interactions with exclusion (see Section 3.3.4) and unweighted by participants.

Two types of interaction metrics were calculated. On the one hand, there are the interaction times calculated from the GameLogs of the *MCI-Exergame*, on the other hand there is the analysed video footage yielding insights regarding the interaction attempts of the study subjects. Although the game is able to distinguish between right and left for all one-sided movements, *boxing R* and *boxing L* as well as *stepping over R* and *stepping over L* were each combined for this usability

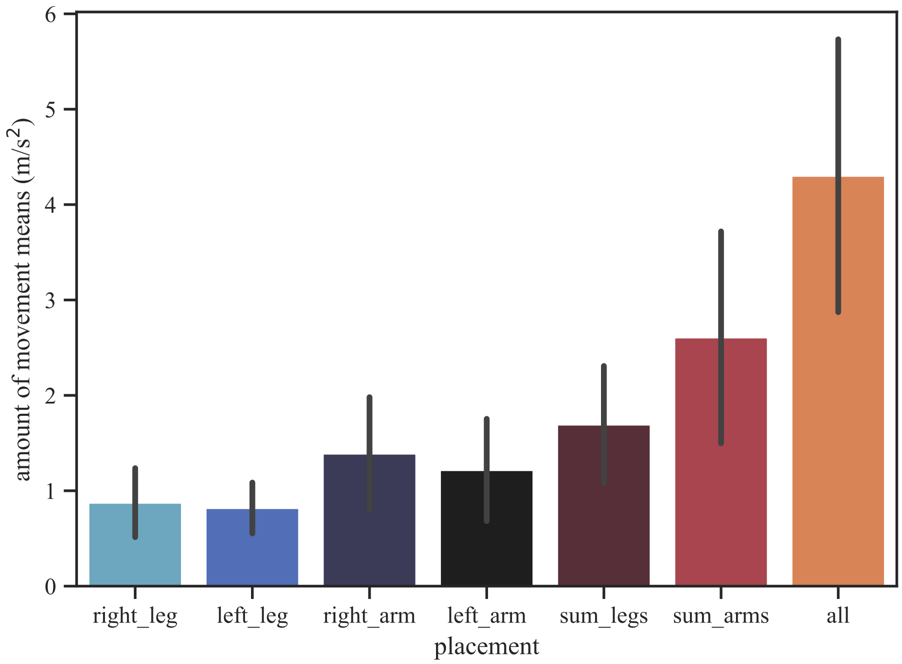


Figure 4.2: Barplot on the evaluation of the amount of movement. The displayed values show the aggregated value over the means of each participant. The black line symbolises the SD between participants and are not based on the whole dataset.

study. This is due to the game implementation. In the current game state both sides of these movements are allowed for interacting with the same object. Interaction metrics were only evaluated for *Level 1*.

For the evaluation of the interaction times (as described in Section 3.3), the three longest times and the 15 shortest times were excluded due to the evaluation criteria (see Section 3.3.4). The resulting analysis shows a sum 1356 interactions for all 17 participants. During the ten minutes (including the connection time), a participant could interact with a mean of 79.79±13.53 obstacles. The average unweighted interaction time across all game objects was 2.206±2.101 s, the mean participant interaction times have an average 2.246±0.457 s regarding the participant means distribution. Table 4.6 shows the obstacle (and resulting movement) distribution from the obstacle point of view unweighted by participants. Additionally, the mean interaction time as well as the number of obstacles which had an interaction time below four seconds are provided. Four seconds, including interaction start, reaction, and detection time, were estimated by the study developers as the amount of time, where the game flow is not impaired yet. Noticeable are the large differences between very high rates for *clapping*, *turning Left* and *turning Right*, and *reaching up R* and *reaching up L* and lower rates for *picking up R* and *stepping over* movements.

movement	count	count one try	count one or two tries	count passed (optional only)	count artificial
reaching up R	102	96 (94.12%)	97 (95.10%)	3 (2.94%)	0 (0.00%)
reaching up L	158	150 (94.94%)	158 (98.73%)	0 (0.00%)	0 (0.00%)
picking up R	98	73 (74.49%)	81 (82.65%)	6 (6.12%)	7 (7.14%)
picking up L	144	102 (70.83%)	122 (84.72%)	13 (9.03%)	3 (2.08%)
boxing	192	180 (93.75%)	186 (96.88%)	0 (0.00%)	0 (0.00%)
clapping	108	106 (98.15%)	107 (99.07%)	1 (0.93%)	0 (0.00%)
jump move	128	113 (88.28%)	127 (99.22%)	0 (0.00%)	0 (0.00%)
stepping over	133	109 (81.95%)	129 (96.99%)	0 (0.00%)	1 (0.75%)
turning left	112	110 (98.21%)	111 (99.11%)	0 (0.00%)	1 (0.89%)
turning right	193	191 (98.96%)	193 (100.00%)	0 (0.00%)	0 (0.00%)

Table 4.7: The table represents the results of the interaction attempts evaluated from the video footage. The numbers presented are based on all performed movements by all participants and are not weighted by participant or movement.

From the evaluated video footage, 17 participants were able to interact with 1368 objects. Here, no data was excluded as the exclusion described above was a *GameLog* specific problem and does not have an influence on the footage evaluation. Table 4.7 gives a overview over the correct interaction attempts per obstacle. As defined in Section 3.3 attempts were counted when they were correctly executed at a point in time when the specific object was showing interactivity. Out of 1368, 1230 (89.91%) interactions took one correct attempt by the participant. Furthermore, for including a second attempt, the number is 1309 (95.69%) out of 1368. For all movements except *picking up*, this results in a proportion greater than 95% of the interactions. *Picking up* also achieved the highest rates for passing objects without interaction. Passing objects was only possible for the optional interactions *picking up* and *reaching up*. *picking up R* required the highest number of artificial detections with 7 (7.14%) out of 98, meaning that the participant had already had three correct attempts without reaching an interaction before the study leader generated an artificial detection.

4.3.3 Questionnaires

With the reverse-coded elements turned the 8-item physical activity enjoyment scale (PACES-8) results in a mean score of 40.82±4.43. On average, a score of 5.10 on a scale of 1-7 was chosen for each statement. Statement 1 - I find it pleasurable was rated particularly positively with 6.06±0.75. Statement 6 - It is very exhilarating was rated the with a minimum average of 4.41±1.18. Figure 4.3 shows a boxplot diagram on the results of the PACES-8. It also visualizes that all medians are

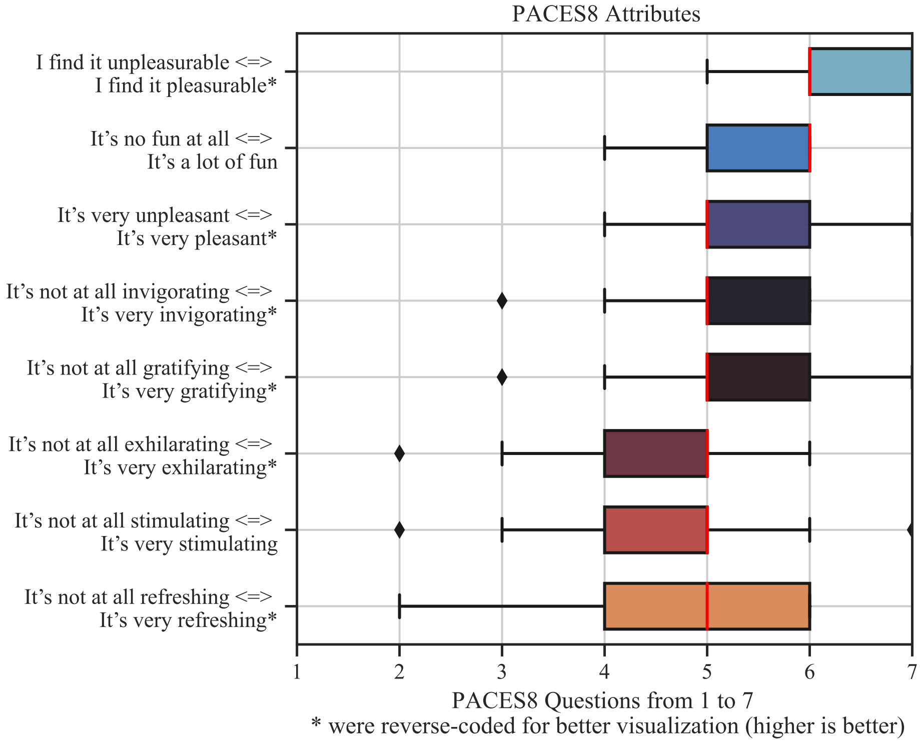


Figure 4.3: Shows a boxplot on the results on the single statements of the PACES-8 from all participants. The median is marked with a red line. Outliers are represented with diamonds

shifted in the direction of the positive alternative.

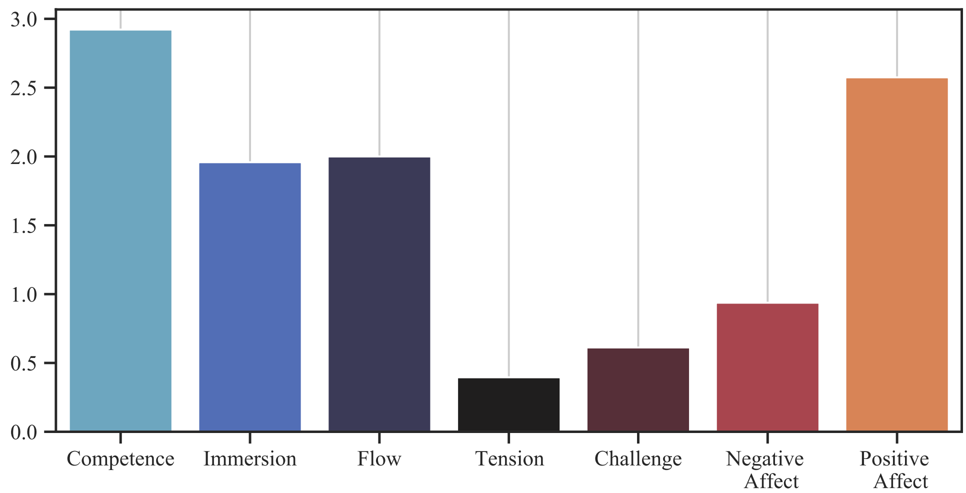


Figure 4.4: A barplot representing the results of the GEQ in the defined categories *Competence, Immersion, Flow, Tension, Challenge, Negative Affect, Positive Affect*.

One participant had to be excluded from the evaluation of the game experience questionnaire (GEQ) because only half of the questionnaire had been answered, resulting in 16 GEQ questionnaires being evaluated. The evaluation of GEQ results in measures in seven categories as a mean value on a scale of 0-4. The results for these categories are the following: *Competence (2.92), Immersion (1.96), Flow (2.00), Tension (0.40), Challenge (0.61), Negative (0.93) and Positive Affect (2.58)*. The results in these categories are visualised in Figure 4.4. The full GEQ boxplot for all questions can be viewed in Appendix D.

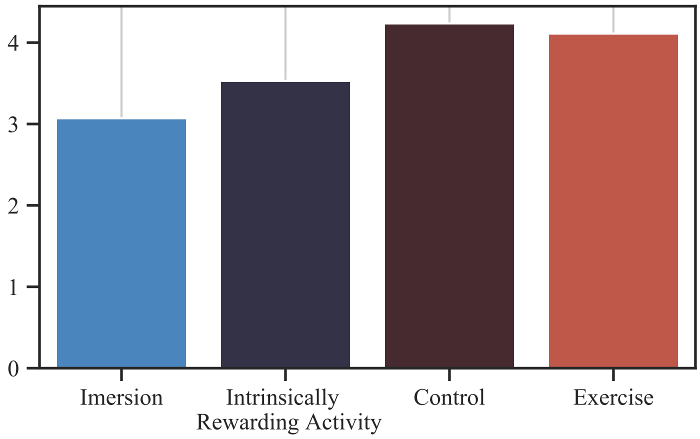


Figure 4.5: Barplot on the results of the EEQ in the defined categories *Immersion, Intrinsically Rewarding Activity, Control, Exercise*.

The exergame experience questionnaire (EEQ) was evaluated for all 17 participants according to the original instruction [Fit20]. The questions were answered on a scale from 1-5. After reverse-coded elements were turned, a summed score for each participant was calculated. The overall EEQ score, which is a mean of these individual scores, is 75.76±5.93. Additionally, scores for the following categories were calculated: *Immersion (3.07), Intrinsically Rewarding Activity (3.53), Control (4.24) and Exercise (4.11).* These scores, which are visualised in Figure 4.5 are also based on a scale of 1-5. The full EEQ boxplot for all questions can be inspected in Appendix D.

The system usability scale (SUS) was completed by 17 participants with a score of 1-5 for each question. The SUS was interpreted according to Brooke et al. [Bro96] resulting in an overall score of 86.47. According to a score interpretation [Ban08] by Bangor et al., this categorizes as the following: Usability is in the top quartile of the tested systems. In terms of acceptability range, it is considered *acceptable*. In the grade system From A to F it is a *B* and in the adjective rating system it accounts for an *excellent*. Figure 4.6 visualizes the answers given in a boxplot.

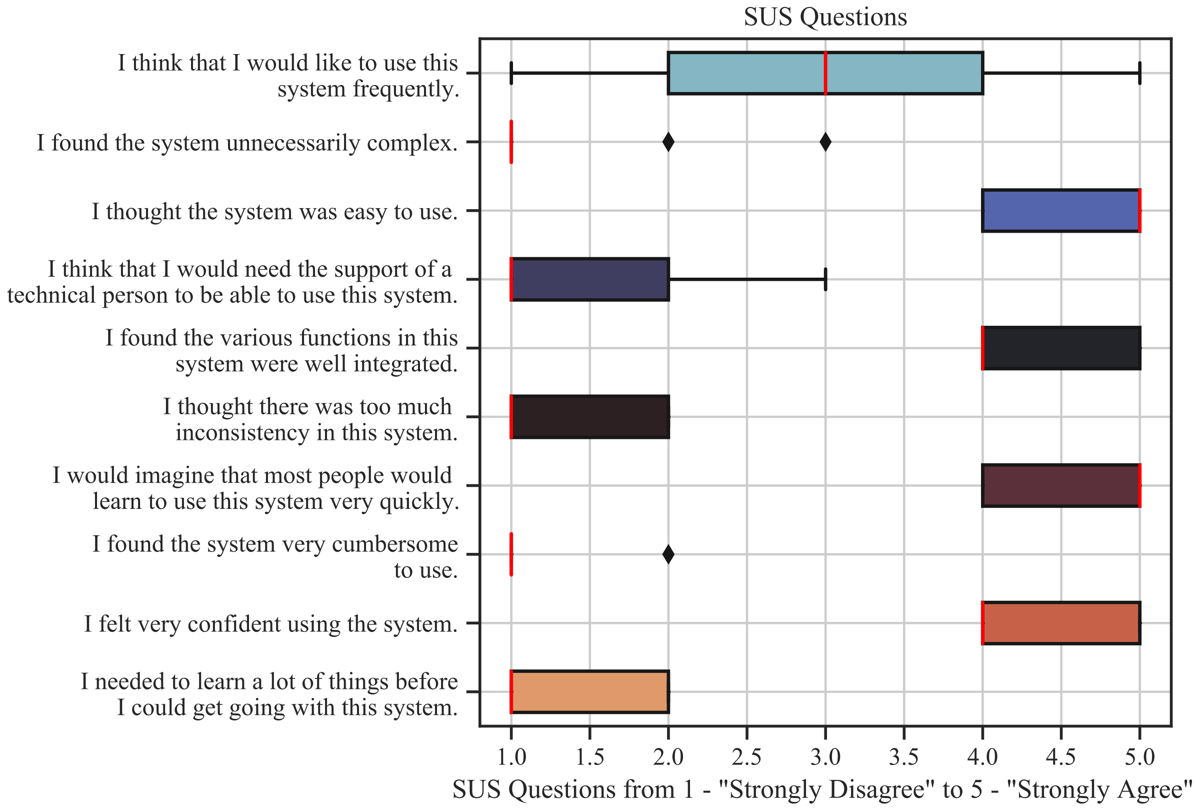


Figure 4.6: Boxplot on the results of the SUS. The median is marked with a red line. Outliers are represented with diamonds

4.3.4 Qualitative Measures

Qualitative results were obtained through the Think aloud method [van94] during the *Tutorial* and *Level 1* and the answers to open questions after the interaction phase of the study.

During the interaction in the tutorial and Level 1, the participants' thoughts were collected using the think-aloud method as described in Section 3.3. A total of 160 statements were recorded and evaluated. All recorded think-aloud statements can be inspected in Appendix D. Seventeen general situational positive statements were recorded and fourteen positive statements directly related to gameplay and design. Four times statements implicating boredom were recorded. Regarding interaction and movement detection. Fourteen statements were recorded that were related to problems with the detection of the *picking up* movements. In addition, eight statements were written that related to problems with the perspective and the resulting timing of the interaction.

All individual answers to the Open Questions can be inspected in Appendix D. Open Questions.

The following is a summary of the answers given for each question:

1. *How long do you think you have played Level 1?*

   The duration of the Level 1 play time was ten minutes including connection time. All participants answered with a numerical value. If a the answer was a range of minutes, the mean was used. The answers of the participants ranged from 3 min to 10 min with a mean of 6.79±2.36 min. No participant estimated the duration to be longer than the completed 10 minutes. One estimation was exactly 10 minutes.

2. *How long would you like to play the game?*

   Before the question was asked, the answer to question 1 was revealed. In contrast to the first question, the participants here mostly answered with a combination of time-based and a qualitative statement. Seven stated an a amount shorter than ten minutes, five wanted to play for ten minutes and five said they wanted to play longer than ten minutes. Six participants additionally argued that they would like to play longer than stated, if there were additional features or levels.

3. *Would you prefer a 2D or 3D game and why?*

   All participants but one stated that they prefer a 3D game environment. One participant stated having no preference. In case of visible uncertainty or when asked, it was explained that 3D game does not mean a virtual reality game. The Participants provided the following reasons for their 3D preference were given: Game interactions seem realer or more immersive (9), three of them related this statement directly to the imagination of the movements), reasons related to excitement or fun (6). One participant stated that 2D might be to abstract for imagining the movements. Another participant found the broad vision and the grass soothing in 3D and again another proposed to change the scenery to the current season.

4. *Did you find the interaction with the game unnecessarily complex and why?*

   For this question, all but one participant stated that they did not find the game unnecessarily complex. One participant did not directly answer the question. Two participants pointed out that *picking up* did not work well. One participant found the standing still for optional objects weird in the beginning, but additionally stated that it was fine later on and people would figure it out. Another found it not as intuitive to turn back towards the screen after *turning* movements. Furthermore, three participants stated explicitly that the movements/interactions were either intuitive or easy to learn. One participant liked that many body parts were used in the game

5. *How many times per week would you play this exergame at home after prior instruction?*

   Nine participants split the answer to this question into how often they would play voluntary and after medical instruction / for medical benefit or after recognizing cognitive decline. Three participants only stated how often they would play voluntarily. One participant only stated the amount in case of medical indication. Three participants stated their answers without any condition. Latter were assigned to both categories. One participant did not state an amount, but that playing regularly would be fine, especially if there were more features. The mean voluntary sessions per week were calculated 1.16±1.27 times out of 15 answers, with six persons stating 0 times and a maximum of 3.5 times stated by one participant. In case of medical indication or use a mean of 5.00±3.35 times was given, ranging from 2 times to 14 times per week.

4.3.5 Correlations

In this study, the Pearson's correlation between SUS on the one hand and interaction metrics on the other hand was evaluated. Pearson's correlation between participant's SUS and mean interaction time for n=17 yielded a r-value of r = 0.1870 and a p-value of p = 0.4967. Interpreted after Cohen [Coh88] this means a nonsignificant mild correlation. Between the system usability score and the percentage of one-try interactions for n=17, a r-value of r = -0.3765, and a p-value of p = 0.1364, meaning a nonsignificant, negative and moderate correlation between both measures. Figure 4.7 visualises the datasets in a scatterplot with a regression line, respectively.

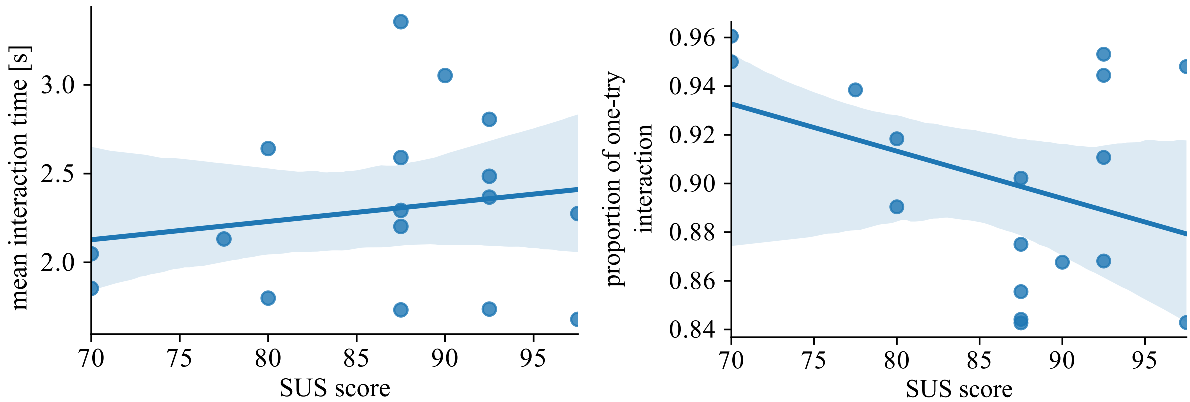


Figure 4.7: The figure presents two scatterplots with regression lines. On the left side SUS score and interaction time are displayed and on the right side SUS score and percentage of one-try interactions are shown. The regression lines only serve to visualise the data. No regression analysis was performed.

Chapter 5

Discussion

The following chapter will first evaluate the already presented results. For this purpose, all parts of the results are discussed in dedicated sections. First, the final state of the exergame is discussed. In particular, the extent to which the objectives have been met, what limitations remain, and possible future developments are assessed. This review of the *MCI-Exergames* contributes strongly to the evaluation of the general feasibility. Afterwards, the results and limitations from the movement detection study are evaluated. The resulting evaluation of the movement detection system helps to answer the first research question regarding the reliability of movement detection using inertial measurement unit (IMU)s (IMU). Then, the outcomes of the usability study are critically discussed. For a better understanding, the usability and enjoyment scores are also compared with other exergames. This assessment contributes to the answer of the second research question regarding the usability of the developed system. Finally, in section 5.4 - Feasability Evaluation, a final assessment of the feasibility is stated and the research questions are answered.

5.1 Exergame Limitations

In this section, the different parts of the *MCI Exergame* development are discussed. In particular, the limitations, improvement possibilities, and possible future work on the exergame are presented.

5.1.1 Data Processing and Movement Detection

Within the scope of this thesis, a new composition of calculation steps for data processing and movement detection was presented. Although both parts were able to prove their general functionality, there are still multiple limitations and a need for further development in many areas.

The initial orientation of the IMUs depending on their placement is currently implemented as an initial guess, which is then adjusted to a high Madgwick β value in the first second to match the global gravity (Section 3.6). This implementation cannot adjust the heading angle around the global axis W_z. While no studies were conducted to determine the resulting deviation from the correct orientation, this could still have an influence on the movement detection. Further work on this topic could include an initial orientation phase consisting of specified movements to estimate the individual orientations more exactly (e.g., pointing towards the TV or device). Another option could be the use of further technology to obtain information on the individual orientation which is continuously updated (e.g., with a magnetometer). With a better orientation, it could also be an option to use the hand (IMU) as a pointer to select items and thereby be fully independent of the touchscreen of the device. This feature has already been implemented in an exergame for older people which also used IMUs [Adc19].

To dedrift velocities and trajectories, Butterworth highpass filters and a complementary filter for accelerometers and barometer were incorporated. While the latter worked fine to stabilize the vertical position of the IMUs in the presented prototype, there were still visible changes in the position due to the inaccuracies of the barometer. Especially, for smaller movements like *stepping over*, other approaches could improve the detection significantly. A zero-velocity update for walking on the spot could be a viable option for this.

Further improvements could include a BLE connection flow optimization to shorten connection times and an improvement of the *GameLog* to ease the evaluation and provide more physical data. Furthermore, future work should concentrate on optimizing the empirically set thresholds in data processing and movement detection and validate the detection system. Additionally, other methods for real-time motion detection for exergame should be investigated. As already introduced in Section 2.4, different machine learning methods could also be a viable approach to detect the players' movements.

5.1.2 Gameplay and Design

To create a working exergame prototype for elderly people with mild cognitive impairment (MCI), a game with interactive objects was developed, which should appropriately challenge the user cognitively as well as physically. In addition, the game should have a high motivation factor and provide a high usability. Within the time frame of the thesis, some functions could not be fully implemented.

Certain components of the game are tailored to specific cognitive domains (sections 3.1.4, 4.1). However, apart from remembering the individual movements, the presented exergame lacks features that are tailored to the area of memory. While there are components in the game that relate to the executive functions area, there are no components in which a decision-making process occurs. In the future, more functions could be added in these areas. Ideas for further development of the game included selecting the direction at intersections by some means (decision-making) and memorising, e.g., figures or animals to be questioned later on in the game (memory).

The current state of the exergame does not provide a story or other motivational factors for the player. With the development of further levels, this could also be addressed. Other motivational factors could include the use of collectable objects such as coins and apples (Section 3.6), to trade in for new features. Furthermore, future studies could focus on identifying the right amount of interactive objects, cognitive tasks, and physical exertion in an exergame for older people with MCI.

Further evaluation of the perceived current state of the exergame by younger players will be discussed in Section 5.3

5.2 Movement Recording Study

The primary aim of the movement recording study was to act as a basis for the development of the MCI exergame. It was intended to provide information about movement execution as well as the difference between instruction sparsity. Thus, the results of the movement recording study had a direct influence on the development of movement detection as a data basis as well as on the gameplay development regarding the amount of necessary movement instruction. Due to the pandemic situation and the absence of further recordings on elderly people, the data from the movement recording study were also used for an initial evaluation of movement detection. Due to the influence of the data on the development of movement detection, the associated results can only be generalised to a limited extent. However, since the movement detection was not directly trained on the recordings, it can represent an initial estimation of the performance.

5.2.1 Movement Validity

Before the movement detection study, the development group composed of the supervisors, the author as well as Linda Becker and Daniel Schöne had a concrete idea of which movements should be incorporated in the game and their execution. Interestingly, a high percentage of recordings showed movements that were executed differently than expected or classified as another movement altogether. The study was divided into three phases with different levels of instruction (phase I - command only, phase II - command with example, phase III - static visual instruction). While there is a slight improvement on the average validity of the movement with a higher level of instruction (phase I - 20.11% to phase III - 18.24%), no instruction level led to an adequate understanding of the intended movements.

An explanation for the low rate of *picking up* in phase II may be due to the verbal example used, where the person is asked to pick up something as if it lays on their nightstand. This could have been a reason why some participants only reached forward instead of down. For the *turning* movement some people only turned their head or a part of their upper body. The lower rates of the *jump move* could be due to its higher complexity. For the *turning* and the *jump move* a decrease in invalid movements was found for phase II and III, suggesting that the higher verbal or visual instruction could let participants perform specific movements more accurately. The results suggest that a even higher level of instruction or training is needed to ensure the correct execution of intended movements. In order to increase the rate of valid movements as well as the autonomy associated with the application, an instructor was developed for the gameplay, which demonstrates the movements beforehand in a *tutorial*. In the future, a movement recording study should take place, where the instructions from the *tutorial* are displayed to validate whether this results in a higher conformity of the correct movements.

5.2.2 Movement Detection

In the following, the results of the final version of the motion detection are evaluated on the recorded movement data. The limitations mentioned above must be taken into account. An aspect that stands out is that all movements reached very high numbers of true negatives. The proportion of false negatives, meaning the proportion that should have been detected but was not, varies greatly depending on the movement (Table 4.1). The same applies to the proportion of false positives. The high number of true negatives can be explained by the fact that many detection processes differ greatly in the detection algorithm, especially if they involve different limbs.

The high number of false positives is also reflected in the precision. While *turning* (R: 88.89%, L: 100%), *clapping* (87.04%) and *stepping sideways* (R: 75.00%, L: 82.61%) reach rather high scores, *stepping over* (R: 60.00%, L: 48.89%) had the lowest precision. A reason for the false positive values and the resulting precision score could be the inclusion of individual movements in others. For example, *stepping over* could be detected when the legs are being lifted for a high step during *turning* or *stepping left/right*. Similarly, a high enough *clapping* or *jump move* might be detected as an additional *reaching up*. This is also suggested by the further evaluation of the false positive distribution (Table 4.3). In both cases, the movement is detected with a change in the vertical position x_z. In this case, a stronger differentiation between the movements might improve these scores significantly. Another option would be an exclusive detection method per extremity, where only one detection of the same limb can be an outcome. Here, a hierarchical approach in the movement detection per limb could be an option.

A reason for the false negatives and the resulting low recall scores for *stepping over* (R: 47.73, L: 40.47) and *picking up* (R: 65.52%, L: 65.62%) could be, that the individual thresholds were set to the wrong amount. Here, one could argue that the exergame should not only classify any movements, but that these should only be recognised if they are executed with a sufficient intensity. In the case of *stepping over*, a threshold of 0.25 m was set. For *picking up* the threshold of 0.45 m was set. In the opinion of the author, these are heights that are also possible for elderly people to reach. Unlike in the game, however, the participants did not receive any feedback in the recording study as to whether the corresponding movement was correct. Possibly the participants would have performed a higher *stepping over* after a first movement in the game which would not be recognised. This raises the question of how far the recognition has to adapt to the player or whether a specific level should be set that can only be reduced according to the physical capabilities of a player.

Furthermore, both movements (*picking up* and *stepping over*) are based on vertical position changes, so that the inaccuracy of the position calculation due to the inclusion of the barometer could play a role as well. For a future recording of movements, the participant could be trained the correct movements beforehand in an individually moderate intensity. These recordings would allow a direct optimization of the underlying parameters of the detection system. Other movements show a high to very high percentage of recall score. As objects are interacted with as soon as the correct movement is detected (even among others), the recall score is considered essential for the playing of the exergame.

While the Hemming loss (described in Section 3.2.4) of 3.40% shows the overall incorrectness of the detections, it is only of limited informative value regarding the performance of movement detection in the game, as it is primarily influenced by a very high number of true negatives. In this sense macro-averaged precision (76.03%), recall (79.46%) and f1-score (77.05%) demonstrate more clearly how the detection performs across all movement classes. Despite the lower level of instruction and the lack of necessary feedback regarding the proper execution of the movements, it can be assumed that the *MCI-Exergame* could have been played with the recorded movements. While the movement detection system can be improved by in many ways and and considering the mentioned limitations of this evaluation, the reliability of a motion detection by IMUs could be shown in an initial evaluation. Due to the limitations of the data basis for this evaluation, this outcome can not be generalised, but needs further studies for validation.

5.3 Usability Study

In this section, the results of the usability study will be discussed. The results included scores and measures on the usability and enjoyment of the developed prototype as well as physical measures and interaction metrics. This part of the work aims to evaluate these results and thus to contribute to the assessment of the feasibility and the answer of the second research question.

Especially for the scores in different usability ratings, it can be useful to offer a basis for comparison. A particularly similar project was evaluated by Adcock et al.[Adc19, Adc20] regarding usability and cognitive training. The project is an exergame for elderly people, which is based on IMUs as well and is part of the Active@Home project. Exergame trainings include tai chi exercises, dancing and cognitive stepping games. In this project the usability was also evaluated with the system usability scale (SUS) and game experience questionnaire (GEQ) which simplifies comparability with the developed *MCI Exergame*. More information on the cognitive outcomes of the Active@Home project can be found in Section 2.2. The 8-item physical activity enjoyment scale (PACES-8) in turn will be compared to usability studies by Bird et [Bir15] and Chu et al[Chu21]. Both exergames are rather different to the developed prototype but can provide a first orientation regarding the outcomes of the usability study.

5.3.1 Physical Measures and Exertion

Overall, the amount of exertion was small. After a short increase, the mean heart rate stayed at a low intensity for the rest of the *level 1*. This was probably due to the young age of the participating players.

Although no direct conclusions can be drawn regarding physical exertion for elderly people, it may be a positive sign that most of the younger participants (24.6±3.0 years) only felt mild exertion. A limitation for this evaluation was that some participants came by bike or brisk walk, so they might have started with different baseline conditions and in particular the course of the heart rate measurement could have been influenced.

5.3.2 Interaction Metrics

For both interaction metrics, interaction time and number of interaction tries, the results from the usability study show very positive results regarding the flow of the game. A mean interaction time, including reaction time of the player, movement execution, and movement detection of 2.206±2.101 s suggests an undisturbed game flow. In terms of average interaction time, *picking up* (R: 3.538±3.609 s, L: 3.609±2.955 s) and *stepping over* (2.501±2.343 s) took particularly long.

A similar situation can be found in the evaluation of the proportions of each movement that only needed one attempt to be recognised. While an average of 89.91% of movements only took one attempt to be recognised, *picking up* (R: 74.49%, L: 70.83%) and *stepping over* (81.95%) had the lowest values. Although the one try percentage cannot be directly compared to the recall values of the movement recording study, both measures should have a high similarity as they evaluate if a movement that was conducted was also recognised. In this sense, especially *stepping over* but also *picking up* increased their relative score. An explanation for this could be the better instruction of the movements through the tutorial and instructor in general. Another option would be that the younger age of the participants led to a more intensive execution of movement. As already stated above, improvements for *stepping over* and *picking up* detection could include a change in the needed vertical height as well as the development of a position calculation with a higher accuracy.

Although there is still room for improvement in the detection of *stepping over* and *picking up*, an overall percentage of 89.91% detections on the first try suggests a reliable interaction with the exergame and the associated movement detection. Future studies should confirm these measures in elderly people. In this context, it could also be shown which contribution the young age of the participants had on the positive results of the study.

5.3.3 Usability and Enjoyment Scores

To evaluate the usability of the developed system, various measurement methods were used to obtain a comprehensive overview. The measurement methods used include the 8-item physical activity enjoyment scale (PACES-8), the exergame experience questionnaire (EEQ), the game experience questionnaire (GEQ) and the system usability scale (SUS). In the following, the results of these scores are discussed and illustrated in comparison with scores from other exergames. The PACES-8 evaluates the physical enjoyment during the activity. From a maximum of 56 points a score of 40.82±4.43 was obtained. There is currently no established interpretation of the PACES-8, other than a higher score suggest a higher enjoyment of physical activity. In comparison to a pilot study by Bird et al. [Bir15] the exergame prototype suggests lower physical enjoyment. In this study 24 elderly people played an off-the-shelf exergame focused on postural balance on the XBOX Kinect. The game training sessions included dancing, cardio-boxing, floor exercises and rope skipping. Here a PACES-8 score of 53.0±0.7 was reached. The great variety of movements and the professional implementation could have contributed to a higher score. In contrast, in the evaluation of a self-developed cognitive stepping game by Chu et al. [Chu21] scores from 39.60 to 46.89 were reached with residential elderly people. This exergame had a lower physical and higher cognitive load for the player. This could suggest, that the complexity and variety of movement would need to be adapted to yield a higher physical activity enjoyment. Another reason for the lower PACES-8 score could be the low physical challenge for the young participants. To further evaluate this, a test with elderly people should be conducted in the future.

As the EEQ is a very new score, references for the score are still missing. The high score of 75.76±5.93 out of 100 might suggest a positive exergame enjoyment. Especially the high value for *control* (mean: 4.24 out of 5) suggests positive feedback for the interaction and the underlying motion detection in general. This is supported by the mean score of 4.53 for the question "I felt that it was easy to familiarize myself with the game controls" and 4.24 for "I felt in control of the game". In contrast, for the statement "I felt that the game reacted quickly to my actions" a score of only 3.29 was achieved. Comparing the score to the original study, it is significantly higher than the evaluated mean score of 67 for Pokemon Go and very similar to the score of 75 for the game Just Dance. As both exergames are quite different from the developed prototype, these scores can only be an orientation.

The evaluation of the GEQ shows a mixed picture. While there were high mean scores (out of 4 points) for the categories *Competence* (2.92) and *positive affect* (2.58) and low scores for the negative categories *negative affect* (0.93) and *tension* (0.40), the categories *immersion* (1.96) and *flow* (2.00) only yielded medium values. As Adcock et al. [Adc19] first calculated the categories for each participant and then took the median of these values, these values were also obtained from this usability study's data for comparison reasons, yielding (median, 1st quartile, 3rd quartile) *competence* (3, 2.5, 3.25), *immersion* (2, 1.63, 2.33), *flow* (2.00, 1.60, 2.5), *tension* (0.17, 0.00, 0.75), *challenge* (0.60, 0.20, 1.00), *negative affect* (0.88, 0.50, 1.13) and *positive affect* (2.60, 2.40, 2.80).

The outcomes for the GEQ in the comparison study by Adcock et al. were: *Competence* (2.3, 2.2, 3.0), *immersion* (2.3, 1.5, 2.7), *flow* (1.0, 0.7, 1.5), *tension* (0.2, 0.0, 0.3)), *challenge* (1.2, 0.7, 1.5), *negative affect* 0.5, 0.2, 0.8) and *positive affect* (2.8, 2.0, 3.3). While the *MCI-Exergame* obtained superior measures in the categories competence and flow, the comparison shows that there is still a need for further development in the direction of immersion and challenge. The differences could be due to the fact that the Active@Home game, unlike the prototype, has a story with several levels. In addition, a higher score in the challenge category might be achieved through different movement sequences and cognitive games. At the same time, it should be noted that the young age of the participants in this usability study could have led to a particularly low challenge.

As the SUS is a reliable source of usability evaluation, it can highly contribute to the feasibility evaluation and the answer to the second research question. The SUS evaluation yielded a high mean score of 86.47. The interpretation of this score has already been part of Section 4.3.3. With a focus on the independent use of the application, it is particularly encouraging that question 3 (3.53 out of 4) "I think that I would need the support of a technical person to be able to use this system"(reverse-coded) and question 10 (3.65 out of 4) "I needed to learn a lot of things before I could get going with this system"(reverse-coded) were rated very positively. For comparison with the Active@Home study, this study's median SUS (87.5, 80, 92,5) was calculated. Regarding the study by Adcock et al. (75.0, 70.0, 85.0) the prototype shows a superior outcome. Limiting factors for this comparison include that the intervention group in the comparison study consisted of elderly people and tested a complete system over a longer duration. For a full comparison, a further developed MCI-Exergame could be tested with elderly people as well.

5.4 Feasability Evaluation

In this master thesis, the feasibility of an exergame for elderly people MCI was investigated on the basis of a self-developed prototype.

From a technological point of view, a reliable prototype was presented. This included a working application prototype for multiple operating systems, a stable and reliable Bluetooth connection with up to 4 IMUs (NilsPods) at the same time, and a working gameplay. It has been shown during this work, that it is feasible to use the 3D development environment Unity in combination with BLE IMUs and to base interactions in the exergame on incoming signals from these sensors. Furthermore, a rudimentary real-time movement detection was presented that relied on basic parameters, including acceleration, velocity, position, and orientation of the 4 sensors. Although the technical part of the prototype withholds the limitations discussed above, the detection system is capable of providing a reliable interaction with the exergame for most movements. This leads to the answer of the first research question placed on this project:

1. Can different movements of old and young be reliably detected with inertial measurement units?

The results from a movement detection with five elderly people as well as the interaction metrics of the usability study suggest a reliable detection of different movements with inertial measurement units. While these results are a positive signal, further well-planned studies with elderly and younger people are needed to confirm this initial indication and to account for the limitations of the studies. In the opinion of the author, it is possible to increase the reliability of the detection system significantly by a further development of data processing and movement detection or by adopting a different method. In the context of the project, it can be presumed that it is possible to develop a movement detection with IMUs, that is able to distinguish movements well enough for use in an exergame.

The developed system was evaluated with an excellent system usability scale score, a high exergame enjoyment score and a moderate game experience questionnaire as well as physical activity enjoyment score. The physical and cognitive challenges in a usability test with younger people were low. While the reaction to some gameplay elements were very positive, multiple components are still missing. These include a storyline, a higher variety of tasks and levels, and further cognitive elements for training against cognitive decline. In the following, the second research question will be answered:

2. How usable is the developed exergame for elderly people with mild cognitive impairment?

The very positive usability results in the second study suggests that the developed exergame is highly usable. Due to the limitations, these results can only be seen as an initial indication. Further studies with elderly people and people with MCI are needed to confirm these measures. As in the second study, enjoyment, cognitive, and physical challenges should also be investigated. To be not only usable but also useful for older people and people with MCI, additional factors should be examined. These include the assessment of injury and fall risks, the ability to use the exergame independently, the adherence rate, and later on, physical and cognitive changes after an extended period of use.

However, from comparison with the literature (Section 2.2), it can be estimated that a game based on simple movements with a low cognitive and physical challenge and a high usability score could be a usable as well as a useful system for older people and people with MCI.

Chapter 6

Conclusion and Outlook

The aim of this thesis was to demonstrate the feasibility of an exergame based on inertial measurement unit (IMU) for people with MCI. For this purpose, an exergame prototype, the *MCI-Exergame*, including a rudimentary movement detection, was developed in Unity. With this exergame prototype, movement detection reliability as well as system usability and enjoyment were evaluated in two studies.

The developed prototype itself operated properly and was able to connect reliably to the four IMUs. The game consisted of a 3D Parkour of interactable objects with limited inclusion of cognitive features. Further developments should include an optimization of data processing and movement detection algorithms as well as more cognitive tasks in the executive functions and memory domain. The gameplay could also be expanded with more objects, a story and several levels, although the cognitive load of patients with mild cognitive impairment (MCI) during the game should be studied beforehand.

The results from the movement detection generally showed positive outcomes, and it was concluded that the detection rates were high enough to base an exergame on the proposed IMU-based detection. While the results are generally encouraging, there are some limitations associated with the recording of the movements, the number of participants, and the different detection rates of individual movements. These limitations are stated in Chapter 5 Discussion at length and should be considered for future studies.

The usability and enjoyment of the developed exergame was rated with an excellent system usability scale (SUS) score, a high exergame experience questionnaire (EEQ) and a moderate game experience questionnaire (GEQ) score. Additionally, the interaction metrics suggest an undisturbed gameplay flow. It was concluded that the developed *MCI-Exergame* is highly usable, but additional studies with a further developed exergame on elderly people are needed to validate these outcomes.

To conclude, this thesis could successfully demonstrate the feasability of an IMU-exergame as well as the movement detection and usability of the developed prototype.

As this master thesis was a feasability study, it could only give a very broad overview on the topic of movement detection with IMUs, gameplay developement for elderly people with MCI and usability evaluation of latter. Future studies should focus on one of these topics to validate and extend the presented findings.

Acronyms

AD Alzheimer's disease.

ADL activities of daily living.

aMCI amnestic mild cognitive impairment.

BDNF brain derived neurotrophic factor.

BLE Bluetooth low energy.

CF complementary filter.

DSM diagnostic and statistical manual of mental disorders.

EEG electroencephalogram.

EEQ exergame experience questionnaire.

GEQ game experience questionnaire.

HR heart rate.

HRV heart rate variability.

IMU inertial measurement unit.

JSON JavaScript Object Notation.

MCI mild cognitive impairment.

MEMS micro-electro-mechanical systems.

MoCA Montreal cognitive assessment.

MRT magnetic resonance tomography.

naMCI non-amnestic mild cognitive impairment.

NCD neurocognitive disorder.

PACES-8 8-item physical activity enjoyment scale.

RCT randomized controlled trial.

RMSSD root mean square of successive RR interval differences.

SUS system usability scale.

TMT trail-making-test.

Appendix A

Related Patents

Interactive physical & cognitive exercise system and method

Publication Number

US20160293033A1

Publication Date

Oct. 6, 2016

Inventor

Cay Anderson-Hanley

Applicant

Cay Anderson-Hanley

Abstract

A system and method for enhancing cognitive function of an individual during exercising use stationary exercise equipment to self-propel the individual's avatar or avatar's point of view through a virtual pathway while interacting to complete a cognitive task including registration of the task, verification of discrimination of basic learning, and performance of manipulation of the cognitive task.

Systems and methods for portable exergaming

Publication Number

US10039981B2

Publication Date

Aug. 7, 2018

Inventor

Brian M. Dugan, Steven M. Santisi, Jean Pierre Latrille, Lieven Nuyttens

Applicant

PEXS LLC

Abstract

In a first aspect, a system for playing a video game is provided that includes (1) one or more sensors adapted to monitor one or more biometric parameters of a user and communicate the one or more monitored biometric parameters (MBPs); (2) a computing device adapted to communicate with the one or more sensors and to receive the one or more communicated MBPs; and (3) a video game having an avatar adapted to move an object on an incline, the video game adapted to execute on the computing device. The video game is adapted to control the avatar to perform an action in the video game based in part on the received one or more communicated MBPs. Numerous other aspects are provided.

Method and system for interactive fitness training program

Publication Number

US8113991B2

Publication Date

Feb. 14, 2012

Inventor

Gershom Kutliroff, Alon Shrut (IL)

Assignee

Omek Interactive, Ltd., Bet Shemesh (IL)

Abstract

A system and method of providing an interactive training program by monitoring a user's actions is disclosed. A user is prompted through the interactive training program to perform certain movements, the user's movements are monitored relative to the prompted movements, and feedback is provided to the user through the program.

Methods and apparatus for a video game magic system

Publication Number

US9545571B2

Publication Date

Jan. 25, 2008

Inventor

Ian Johnson; Dominic Jackson; Nouman Hanif

Assignee

Nintendo Co., Ltd., Kyoto (JP)

Abstract

A system for allowing a player to invoke magic or other special powers in a video game is provided. To activate the magic, the player moves a motion detecting controller in accordance with one or more provided instructions. The outcome of the activation of the magic can be determinate on, for example, speed, accuracy, etc. Controllers in one or both hands may be used, and the instructions can be as simple as single direction gestures and as complex as multi-directional symbols which must be traced in the air. A sequence of any type of instructions may also be provided to instruct the activation of the magic.

Appendix B

Movement Recording Study Materials

Movement Recording Study Protocol

Study Protocol

Scientific background

Patients with mild cognitive impairment suffer from cognitive degeneration, which can eventually lead to Alzheimer's disease. Physical and cognitive stimulation through exergames could be an effective way to slow down or even prevent the progression of this condition. Exergames based on Wearable Sensors (Inertial Measurement Units) could enable patients to train these physical and cognitive activities at home on a regular basis.

Therefore, the goal of this study is to record specified aerobic movements of healthy elderly people with IMUs to develop an exergame based on these movements.

Study setup and technical requirements

- All movement recordings will be taken on 02.11. and 09.11.2020 at Henkestrasse 91, 91052 Erlangen.
- A video camera will be used to record the movement sequences.
- A projector will be used to show the slides of recording phase III (see *10.*)
- Four 6DOF IMUs (Figure 1) from Portabiles GmbH will be used for movement recordings. One IMU will be fixated on each wrist with an elastic watchstrap. The other two IMUs will be fixated on the upper part of each shoe with a clip. The IMUs will be oriented in a way that the lower part of the "P" on the housing points towards the back of the hand or the tip of the foot.
- A tablet or smartphone with Android operating system will be used to record the streamed data with the PortabilesDemoApp using Low Energy Bluetooth.

Figure 1 - Picture of NilsPod IMU
(Source: www.portabiles.de)

Preparation

- The IMUs, HRV sensors, iPad and video camera will be checked for functionality and battery charge.

Test sequence

A test sequence will take about 40 – 60 min and will include the following steps:

1. The study participant will be explained the study process and asked to fill in the following documents: Profile of Subject, Physical Activity Readiness Questionnaire (PAR-Q), Declaration of consent.
2. The PortabilesDemoApp will be started and connected with all four sensors.
3. The four sensors will be attached to the corresponding body parts of the participant (see *study setup and technical requirements*).
4. The supervisor will note which IMU is fixated on which body part.
5. The participant will be explained the initial posture. The initial posture consists of the participant should stand on both feet and leave his/her arms hanging.
6. The participant will be asked to stand in the initial posture for the first 5 seconds of the recording and then perform the movement for 30 seconds:
    a. Walking on spot with both feet
7. The participant will be randomly assigned to either *instruction group I* or *II*.
8. The participant will be asked to perform the following movements in three phases (see *9., 10.*):
    1. Reaching up with right arm
        i. General instruction: Reach up with your right arm.
        ii. Specific instruction: Reach up with your right arm to get an apple hanging above your head.
        iii. Visual instruction: Slide with picture and general instruction.
    2. Reaching up with left arm
        i. General instruction: Reach up with your left arm.
        ii. Specific instruction: Reach up with your left arm to get an apple hanging above your head.
        iii. Visual instruction: Slide with picture and general instruction.
    3. Reaching down with right arm
        i. General instruction: Reach down with your right arm.
        ii. Specific instruction: Pick up your wallet with your right arm from the nightstand in front of you.
        iii. Visual instruction: Slide with picture and general instruction.
    4. Reaching down with left arm
        i. General instruction: Reach down with your left arm.

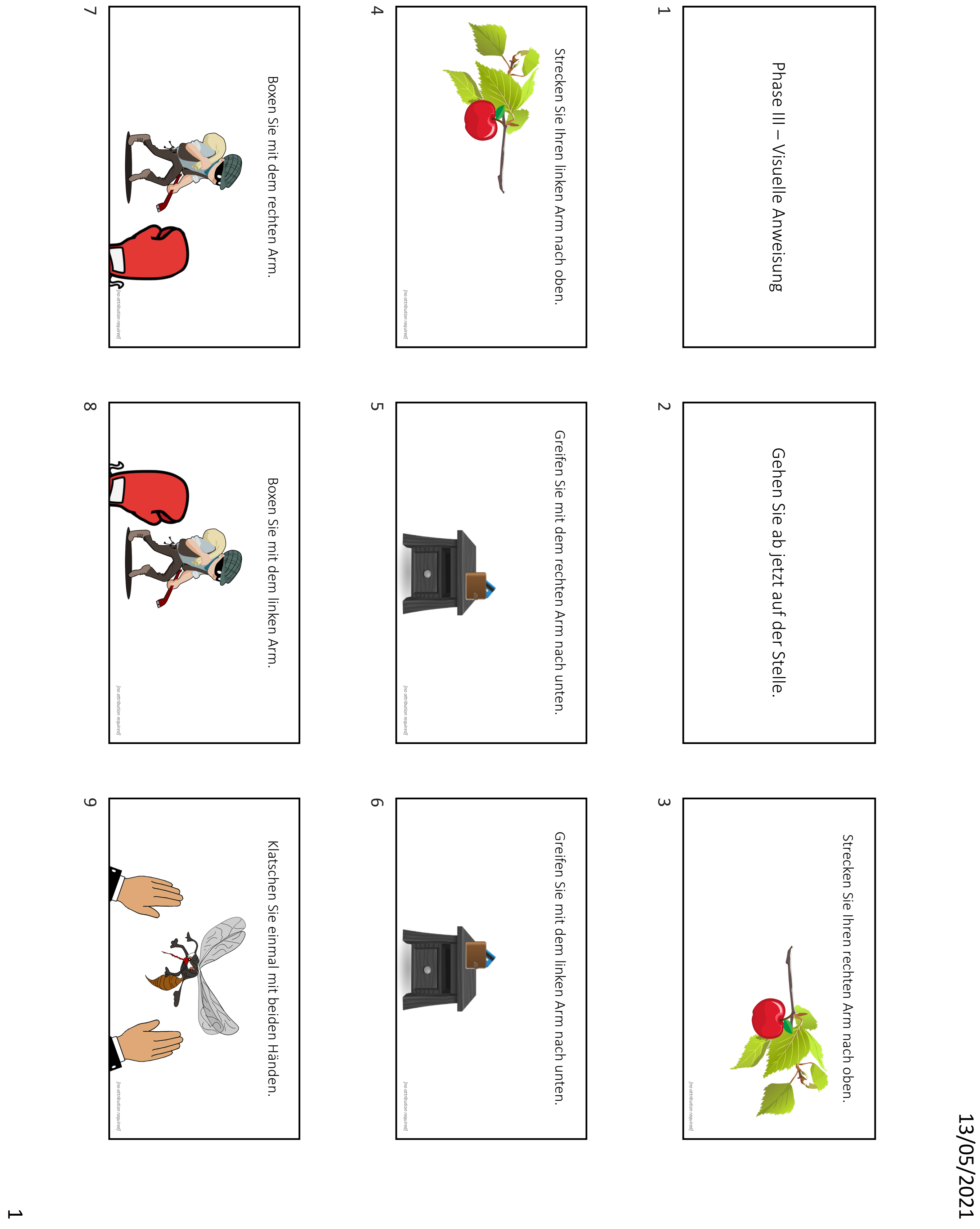
   ii. Specific instruction: Pick up your wallet with your left arm from the nightstand in front of you.
   iii. Visual instruction: Slide with picture and general instruction.
5. Boxing with right arm
   i. General instruction: Box with your right arm.
   ii. Specific Instruction: Box a bank robber right in front of you with your right arm.
   iii. Visual instruction: Slide with picture and general instruction.
6. Boxing with left arm
   i. General instruction: Box with your left arm.
   ii. Specific Instruction: Box a bank robber right in front of you with your left arm.
   iii. Visual instruction: Slide with picture and general instruction.
7. Clapping with both hands
   i. General instruction: Clap with both hands once.
   ii. Specific instruction: Clap to kill a mosquito.
   iii. Visual instruction: Slide with picture and general instruction.
8. Jumping off without losing contact with the ground.
   i. General instruction: Make a jump move without losing contact with the ground.
   ii. Specific Instruction: Make a jump move to get over a little river without losing contact with the ground.
   iii. Visual instruction: Slide with picture and general instruction.
9. Stepping over on spot with the right foot first
   i. General instruction: Step over on spot with the right foot first.
   ii. Step over a branch on spot with the right foot first.
   iii. Visual instruction: Slide with picture and general instruction.
10. Stepping over on spot with left foot first
   i. General instruction: Step over on spot with the left foot first.
   ii. Step over a branch on spot with the left foot first.
   iii. Visual instruction: Slide with picture and general instruction.
11. Turning right and back
   i. General instruction: Turn right and then back.
   ii. Specific instruction: Turn right at an intersection and turn back afterwards.
   iii. Visual instruction: Slide with picture and general instruction.
12. Turning left and back
   i. General instruction: Turn left and then back.
   ii. Specific instruction: Turn left at an intersection and turn back afterwards.
   iii. Visual instruction: Slide with picture and general instruction.
13. Stepping quickly to the right with both feet

    i. General instruction: Step quickly to the right with both feet.
    ii. Specific instruction: Step to the right to avoid stepping in dog poo.
    iii. Visual instruction: Slide with picture and general instruction.

14. Stepping quickly to the left with both feet
    i. General instruction: Step quickly to the left with both feet.
    ii. Specific instruction: Step to the left to avoid stepping in dog poo.
    iii. Visual instruction: Slide with picture and general instruction.

9. If the participant was assigned to *instruction group I*, he or she will be given i) general instruction for each movement in the first phase. In the second phase, the participant will be given the ii) specific instruction. If the participant is assigned to *instruction group II*, he or she will start with ii) specific instruction for the movement in phase I. Afterwards he or she will be given the i) *general instruction* in phase II. The movement order for phase I and II are randomized for each participant. The participant should take care to perform the movement in place and as naturally as possible. The supervisor does not perform the movements and does not give feedback to the participant.
10. For all movement sequences in phases I and II, the participant will be asked to stand in the initial posture in the first 5 seconds of the recording, then he or she will perform the movement on each start signal from the supervisor and return to the initial posture afterwards. The next start signal will be 5 seconds after the return to the initial position. Each movement in phase I and II will be performed 5 times by the participant.
11. In phase III the participant will be shown iii) visual instruction of the movements in *8.*, each consisting of a picture for visualization and the general instruction. Each movement in *8.* will be part of this sequence at least once. The movement order in phase III will be randomised for each participant. In phase III the participant will be also be asked to stand in the initial position in the first 5 seconds of the recording. When the slideshow starts the participant will be asked to walk on the spot continuously as long as this is compatible with the movement. Each time a new slide appears the participant will be asked to perform the movement once.
12. After each movement sequence, the data will be saved from the ProtabilesDemoApp to the internal storage of the smartphone or tablet. In phase I and II this occurs after 5 movements, when a change in movement takes place. In phase III this is after the complete sequence. In this time the participant will be asked to sit down for at least 30 seconds.
13. After phases III is recorded and saved the video camera will be stopped and the IMUs disconnected from the PortabilesDemoApp.

Movement Recording Phase III



Machen Sie eine Sprungbewegung, ohne dass Ihre Füße den Boden verlassen.

[original: www.freevector.com]

10

Steigen Sie auf der Stelle über etwas hinüber.
Starten Sie mit dem rechten Fuß.

rd]

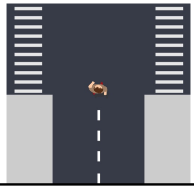
11

Steigen Sie auf der Stelle über etwas hinüber.
Starten Sie mit dem linken Fuß.

[no attribution required]

12

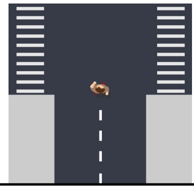
Drehen Sie sich nach rechts und dann wieder zurück.



[no attribution required]

13

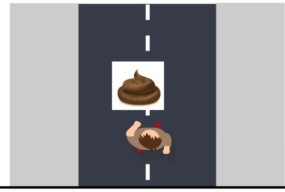
Drehen Sie sich nach links und dann wieder zurück.



[no attribution required]

14

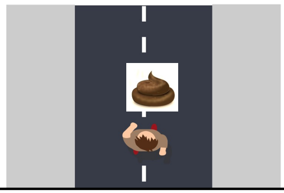
Gehen Sie schnell einen Schritt nach rechts.



[no attribution required]

15

Gehen Sie schnell einen Schritt nach links.



[no attribution required]

16

Appendix C

Usability Study Materials

Usability Study Protocol

Study Protocol

Scientific background

Patients with mild cognitive impairment suffer from cognitive degeneration, which can eventually lead to Alzheimer's disease. Physical and cognitive stimulation through exergames could be an effective way to slow down or even prevent the progression of this condition. Exergames based on Wearable Sensors (Inertial Measurement Units) could enable patients to train these physical and cognitive activities at home on a regular basis.

For this purpose, an exergame was developed as part of a master's thesis that can be controlled by the players' movements in the real world. The objective of this particular study is to evaluate the usability of the developed system as well as the gaming experience and the enjoyment of the physical activity of the exergame.

Study setup and technical requirements

- The study runs will take place from 09/04/2021 to 14/04/2021 at the Machine Learning and Data Analytics Lab, Department of Computer Science, Carl-Thiersch-Straße 2b, 91052 Erlangen.
- Four 6DOF IMUs (Figure 1) from Portabiles GmbH will be used for controlling the exergame. One IMU will be fixated on each wrist with an elastic watch strap. The other two IMUs will be fixated on the upper part of each shoe with a fixation clip. The IMUs will be oriented in a way that the arrow points towards the hand or end of the foot respectively.
- An iPad (5th generation / 2017) with iOS 14.4 and preinstalled *MCI_Exergame App* will be used to play the exergame and the IMUs data will be streamed using Low Energy Bluetooth.
- The iPad screen will be mirrored on a TV screen or a projector using an HDMI Cable and Adapter or streaming via AppleTV.
- A Biovotion Sensor will be placed on the upper right arm to record the heart rate variability of the participant.
- A video camera will be used to record the participant and the TV or projector screen.

Figure 1 - Picture of NilsPod IMU
(Source: www.portabiles.de)

Measurements

Objective Measurements:

- Application Logs will automatically be saved locally in the iPad storage in CSV format at the end of the game and every 2 min after the game started. These logs include:
    1. IMU Data logs for each pod with all streamed values since all IMUs were subscribed in a pre-processed form (same format as recordings from *PortabilesDemoApp*).
    2. A detection log, which includes all detection runs from the movement detection system. In contrast to the normal game mode, the detection system has been changed, so all movement detections run all the time in this study.
    3. A game log, which includes all events in the game, e.g. game start, level loaded, interaction start, and end with an obstacle. Step detections are also included here as they are processed in a separate detection system and currently cannot be saved in the detection log.
- In retrospect, by recording the IMU data the amount of movement can be calculated.
- Heart Rate Variability is measured and recorded using a Biovotion Sensor.

Subjective Measurements

- Questionnaires:
    1. SUS – System Usability Scale (SAP UX Community German translation)
    2. GEQ – Game Experience Questionnaire Core Module (German translation)
    3. EEQ – Exergame Enjoyment Questionnaire (German Translation)
    4. PACES-8 – Physical Activity Enjoyment Questionnaire (German translation)
    5. Borg Scale of Perceived Exertion (German translation)
- Open Questions:
    1. How long do you think you have played Level 1?
    2. How long would you like to play the game?
    3. Would you prefer a 2D or 3D game and why?
    4. Did you find the interaction with the game unnecessarily complex and why?
    5. How many times per week would you play this exergame at home after prior instruction?
- Think aloud – The Participant should say out loud any thoughts about the game or situation during playtime.

Preparation

- The IMUs, HRV sensor, iPad, and video camera will be checked for functionality and battery charge.
- All existing log files will be transferred to a computer for backup.

Test sequence

1. The study participant will be explained the study process and asked to fill in the following documents: Profile of Subject, Physical Activity Readiness Questionnaire (PAR-Q), Declaration of consent.
2. The HRV sensor and the IMUs will be attached to the respective parts of the body.
3. The participants will be given the following instructions for the tutorial phase:
    a. Say out loud any thoughts about the game or situation while playing.
    b. Any movement which the participant needs to do is displayed on the monitor. Please watch the whole instruction before attempting the movement.
    c. A game object is interactable as soon as it is changing its color to light blue and staying interactable until the color changed back to the original.
    d. There are objects you need to interact in the game to move forward and others where interaction is optional.
    e. If a movement is not detected, wait approximately 3 seconds and then try it again. If the movement was not detected after 3 times, the study instructor will generate the movement artificially.
    f. Play the game until the game instructs you to stop. Please stand still after the game has instructed you to stop.
4. The IMUs will be selected in the menu and the tutorial will be set up for the participant.
5. The participant will play the tutorial.
6. As soon as the tutorial ended the log files will be saved to the local storage.
7. The participant will be explained the Borg scale.
8. The participants will be given the following instructions for the level 1 phase:
    a. Say out loud any thoughts about the game or situation while playing.
    b. In level 1, the instructions only appear when you need them. If you don't know what to do stand still until the instruction appears.
    c. A game object is interactable as soon as it is changing its color to light blue and staying interactable until the color changed back to the original.
    d. If a movement is not detected, wait approximately 3 seconds and then try it again. If the movement was not detected after 3 times, the study instructor will generate the movement artificially.
    e. Play the game until the game instructs you to stop. Please stand still after the game has instructed you to stop.
9. The IMUs will be selected in the menu and level 1 will be set up for the participant.
10. The participant will play level 1.
11. During level 1, the participant will be asked for the Borg Scale of Perceived Exertion every 3 minutes.
12. As soon as the tutorial ended the log files will be saved to the local storage.
13. After finishing level 1, the HRV sensor and the IMUs are detached from the participant.
14. The participant will be asked to fill in the following questionnaires: the GEQ, the SUS, PACES-8, and EEQ and answer the open questions.

German Translation of GEQ Core

Game Experience Questionnaire – Core Module

Fragebogen zum Erleben des Games (1)

Bitte geben Sie auf der folgenden Skala für jedes Item an, wie Sie sich fühlten, während Sie das game spielten.

	Gar nicht	Eher nicht	Teils/ teils	Eher stark	Sehr stark
1. Ich fühlte mich zufriedengestellt.					
2. Ich fühlte mich geschickt.					
3. Ich war an der Geschichte des Games interessiert.					
4. Ich dachte, dass das Game Spaß machte.					
5. Ich war ganz eingenommen von dem Game.					
6. Ich fühlte mich glücklich.					
7. Es verschaffte mir eine schlechte Stimmung.					
8. Ich dachte an andere Dinge.					
9. Ich fand es ermüdend.					
10. Ich fühlte mich kompetent.					
11. Ich fand es schwierig.					
12. Es war ästhetisch ansprechend.					
13. Ich habe alles um mich herum vergessen.					
14. Ich fühlte mich gut.					
15. Ich war gut darin.					
16. Ich war gelangweilt.					

German Translation of GEQ Core

17. Ich hatte das Gefühl, erfolgreich zu sein.					
18. Ich hatte das Gefühl, ideenreich zu sein.					
19. Ich hatte das Gefühl, dass ich Dinge auskundschaften konnte.					
20. Ich hatte Freude daran.					
21. Ich hatte das Gefühl, die Ziele des Games schnell zu erreichen.					
22. Ich war verärgert.					
23. Ich fühlte mich unter Druck					
24. Ich war empfindlich.					
25. Ich habe das Zeitgefühl verloren.					
26. Ich fühlte mich gefordert.					
27. Ich fand es eindrucksvoll.					
28. Ich war sehr auf das Game konzentriert.					
29. Ich war frustriert.					
30. Es fühlte sich wie eine große Erfahrung an.					
31. Ich habe die Verbindung zur Außenwelt verloren.					
32. Ich hatte das Gefühl, unter Zeitdruck zu stehen.					
33. Ich musste mir viel Mühe geben.					

17. Ich hatte das Gefühl, erfolgreich zu sein.					
18. Ich hatte das Gefühl, ideenreich zu sein.					
19. Ich hatte das Gefühl, dass ich Dinge auskundschaften konnte.					
20. Ich hatte Freude daran.					
21. Ich hatte das Gefühl, die Ziele des Games schnell zu erreichen.					
22. Ich war verärgert.					
23. Ich fühlte mich unter Druck					
24. Ich war empfindlich.					
25. Ich habe das Zeitgefühl verloren.					
26. Ich fühlte mich gefordert.					
27. Ich fand es eindrucksvoll.					
28. Ich war sehr auf das Game konzentriert.					
29. Ich war frustriert.					
30. Es fühlte sich wie eine große Erfahrung an.					
31. Ich habe die Verbindung zur Außenwelt verloren.					
32. Ich hatte das Gefühl, unter Zeitdruck zu stehen.					
33. Ich musste mir viel Mühe geben.					

German Translation of EEQ

The Exergame Enjoyment Questionnaire (EEQ)

Fragebogen zum Vergnügen beim Exergaming

	stimme überhaupt nicht zu	stimme nicht zu	neutral	stimme zu	stimme voll und ganz zu
1. Ich war angeregt von den physischen Aktivitäten im Game.					
2. Die körperliche Anstrengung im Game gab mir ein gutes Gefühl.					
3. Ich hatte das Gefühl, das Zeitgefühl während dem Spielen zu verlieren.					
4. Ich hatte das Gefühl, zu verstehen wie das Game funktioniert.					
5. Ich habe mich auf das Game konzentriert.					
6. Ich hatte das Gefühl, das Game sei angenehmer ohne die physische Aktivität.					
7. Ich hatte das das Gefühl, es war leicht sich mit der Steuerung des Games vertraut zu machen.					
8. Ich fühlte mich emotional verbunden mit dem Game.					
9. Ich ziehe in Erwägung das Game als „Training“ zu spielen.					
10. Ich hatte das Gefühl, die physische Aktivität sei zu intensiv für mich.					
11. Ich hatte nicht das Verlangen, in dem Game Fortschritte zu machen.					
12. Ich hatte ein starkes Gefühl, in der Welt des Games zu sein, bis zu dem Punkt, dass ich meine Umgebung nicht mehr wahrnahm.					

13. Ich würde lieber nicht Sport machen, auch wenn die körperliche Anstrengung verbunden war mit Elementen eines Games.					
14. Ich hatte das Gefühl, das Game sei gut für mein physisches Wohlergehen.					
15. Ich hatte das Gefühl, das Game bot eine angenehme Herausforderung.					
16. Ich hatte vom Spielen des Games ein Gefühl von Erfolg.					
17. Ich hatte das Gefühl, das Game reagierte schnell auf meine Handlungen.					
18. Ich verspürte nicht den Wunsch weiter zu spielen.					
19. Ich würde es bevorzugen, wenn diese physische Aktivität nicht verbunden wäre mit Elementen eines Games.					
20. Ich hatte das Gefühl, das Game unter Kontrolle zu haben.					

German Translation of SUS Questionnaire

Fragebogen zur System-Gebrauchstauglichkeit

1. Ich denke, dass ich das System gerne häufig benutzen würde.

Stimme überhaupt nicht zu 1	2	3	4	Stimme voll zu 5
○	○	○	○	○

2. Ich fand das System unnötig komplex.

Stimme überhaupt nicht zu 1	2	3	4	Stimme voll zu 5
○	○	○	○	○

3. Ich fand das System einfach zu benutzen.

Stimme überhaupt nicht zu 1	2	3	4	Stimme voll zu 5
○	○	○	○	○

4. Ich glaube, ich würde die Hilfe einer technisch versierten Person benötigen, um das System benutzen zu können.

Stimme überhaupt nicht zu 1	2	3	4	Stimme voll zu 5
○	○	○	○	○

5. Ich fand, die verschiedenen Funktionen in diesem System waren gut integriert.

Stimme überhaupt nicht zu 1	2	3	4	Stimme voll zu 5
○	○	○	○	○

6. Ich denke, das System enthielt zu viele Inkonsistenzen.

Stimme überhaupt nicht zu 1	2	3	4	Stimme voll zu 5
○	○	○	○	○

7. Ich kann mir vorstellen, dass die meisten Menschen den Umgang mit diesem System sehr schnell lernen.

Stimme überhaupt nicht zu 1	2	3	4	Stimme voll zu 5
○	○	○	○	○

8. Ich fand das System sehr umständlich zu nutzen.

Stimme überhaupt nicht zu 1	2	3	4	Stimme voll zu 5
○	○	○	○	○

9. Ich fühlte mich bei der Benutzung des Systems sehr sicher.

Stimme überhaupt nicht zu 1	2	3	4	Stimme voll zu 5
○	○	○	○	○

10. Ich musste eine Menge lernen, bevor ich anfangen konnte das System zu verwenden.

Stimme überhaupt nicht zu 1	2	3	4	Stimme voll zu 5
○	○	○	○	○

8-item PACES Questionnaire

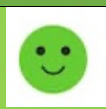
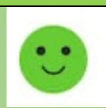
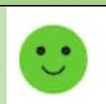
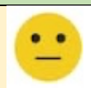
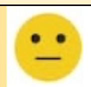
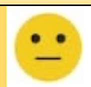
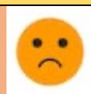
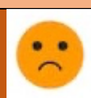
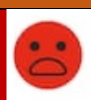
Date ___ ___/___ ___/___ ___ ___ ___ ExaminerID ___ ___ ID: ___ ___ ___ ___ ___ ___ ___ ___

Physical Activity Enjoyment Scale for older adults (PACES-8)

Ich finde es angenehm	❑	❑	❑	❑	❑	❑	❑	Ich finde es unangenehm
Es macht überhaupt keinen Spaß	❑	❑	❑	❑	❑	❑	❑	Es macht viel Spaß
Es ist sehr wohltuend	❑	❑	❑	❑	❑	❑	❑	Es ist überhaupt nicht wohltuend
Es ist sehr belebend	❑	❑	❑	❑	❑	❑	❑	Es ist überhaupt nicht belebend
Es ist sehr befriedigend	❑	❑	❑	❑	❑	❑	❑	Es ist überhaupt nicht befriedigend
Es ist sehr anregend	❑	❑	❑	❑	❑	❑	❑	Es ist überhaupt nicht anregend
Es ist überhaupt nicht stimulierend	❑	❑	❑	❑	❑	❑	❑	Es ist sehr stimulierend
Es ist sehr erfrischend	❑	❑	❑	❑	❑	❑	❑	Es ist überhaupt nicht erfrischend

German Translation of Borg Scale

Borg-CR10-Skala

0	Ruhe	
1	Sehr leicht	
2	Leicht	
3	Mäßig	
4	Etwas anstrengend	
5	Anstrengend	
6		
7	Sehr anstrengend	
8		
9	Sehr, sehr anstrengend	
10	Wie mein härtester Wettkampf	

Open Questions

Open Questions:

1. How long do you think you have played Level 1?

2. How long would you like to play the game?

3. Would you prefer a 2D or 3D game and why?

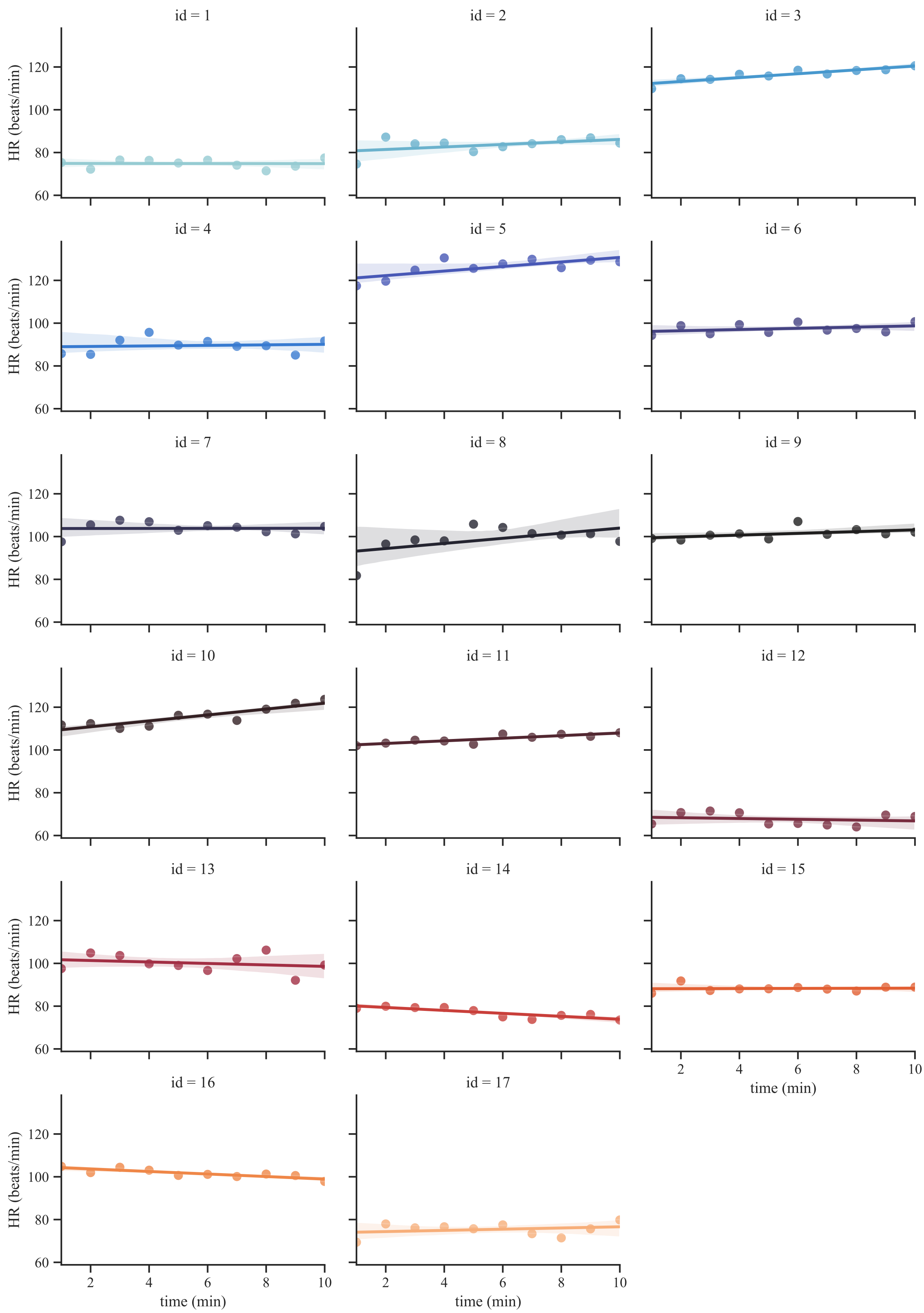
4. Did you find the interaction with the game unnecessarily complex and why?

5. How many times per week would you play this exergame at home after prior instruction?

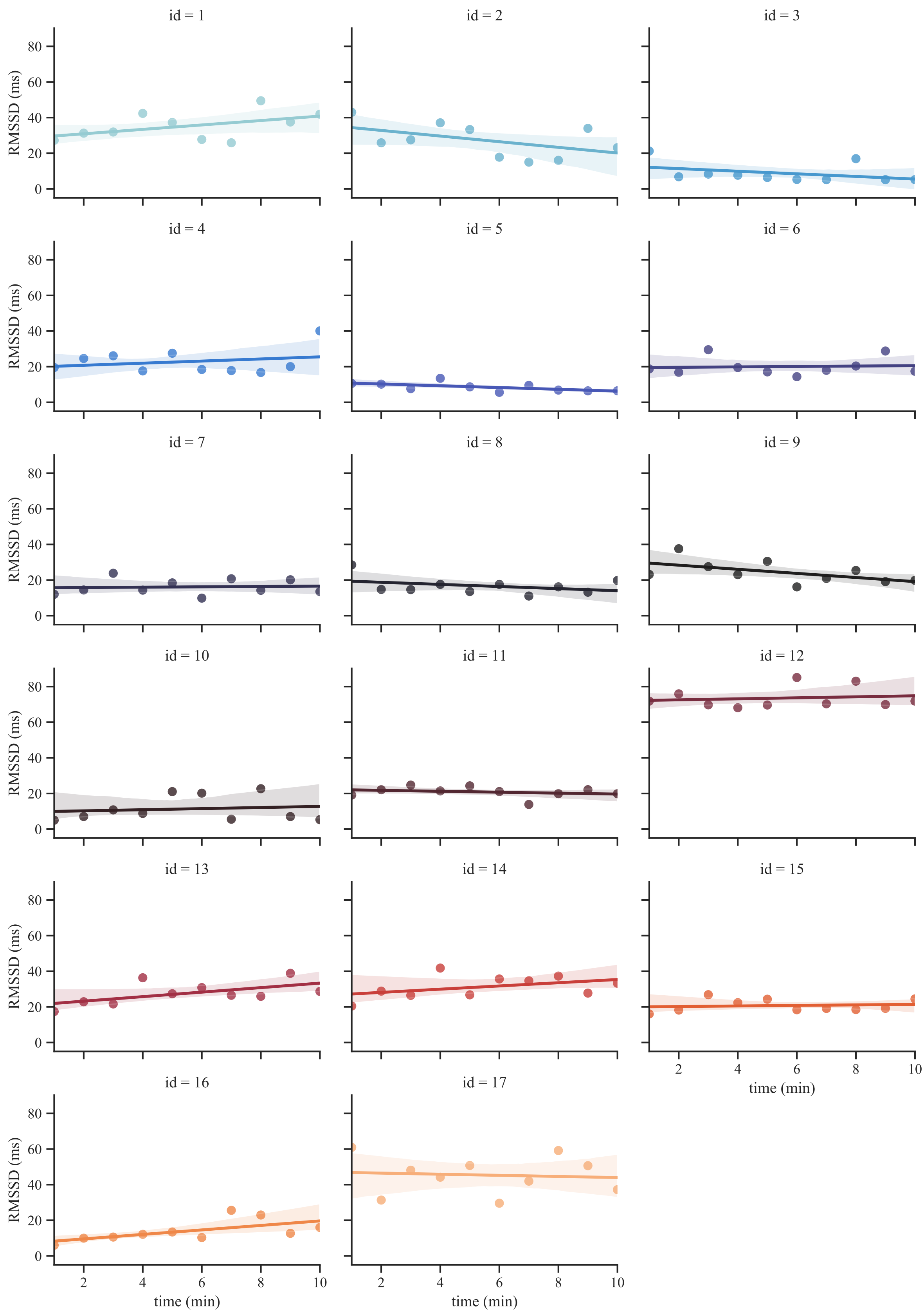
Appendix D

Additional Usability Study Results

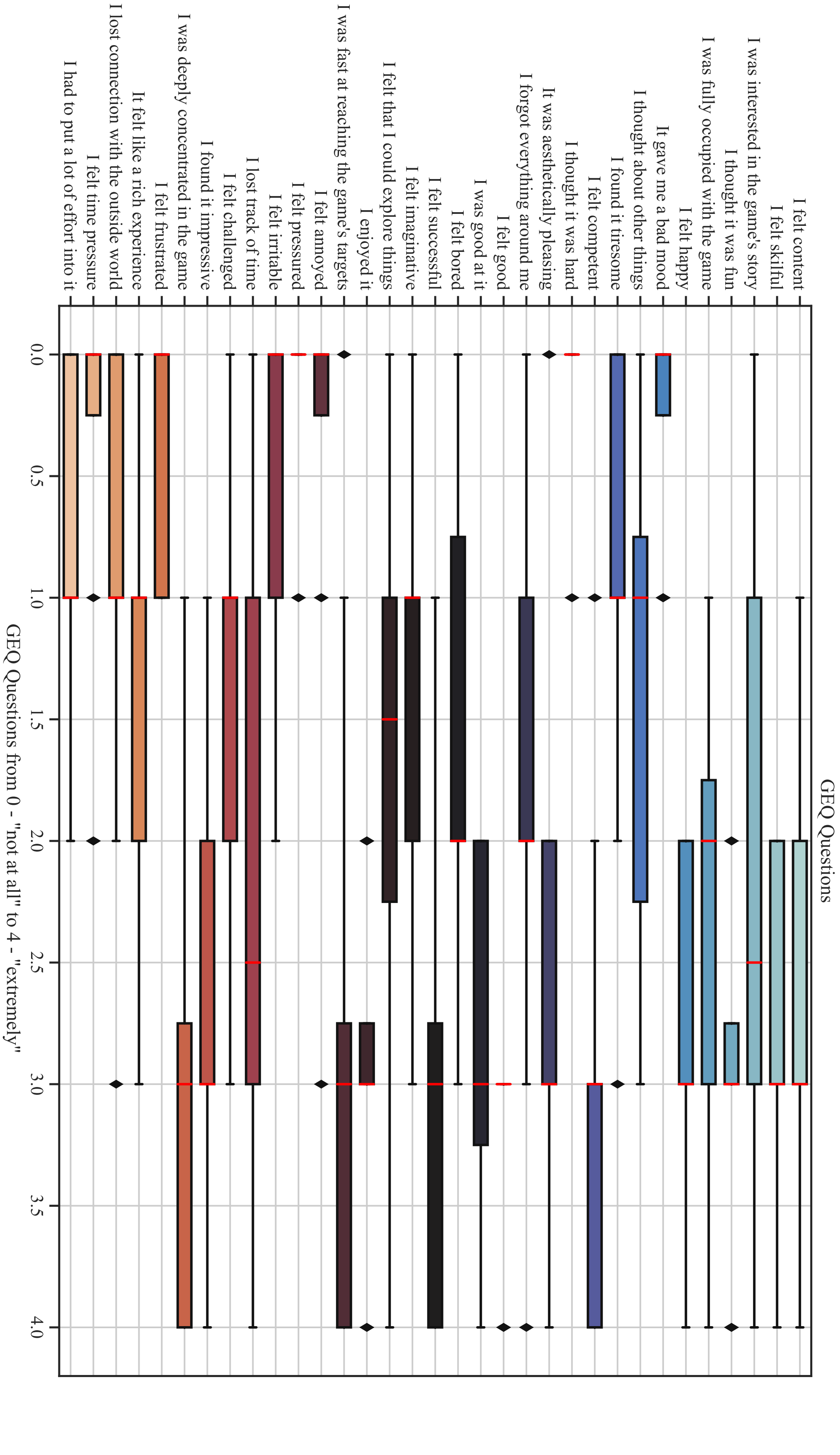
HR Participants Level 1 plots



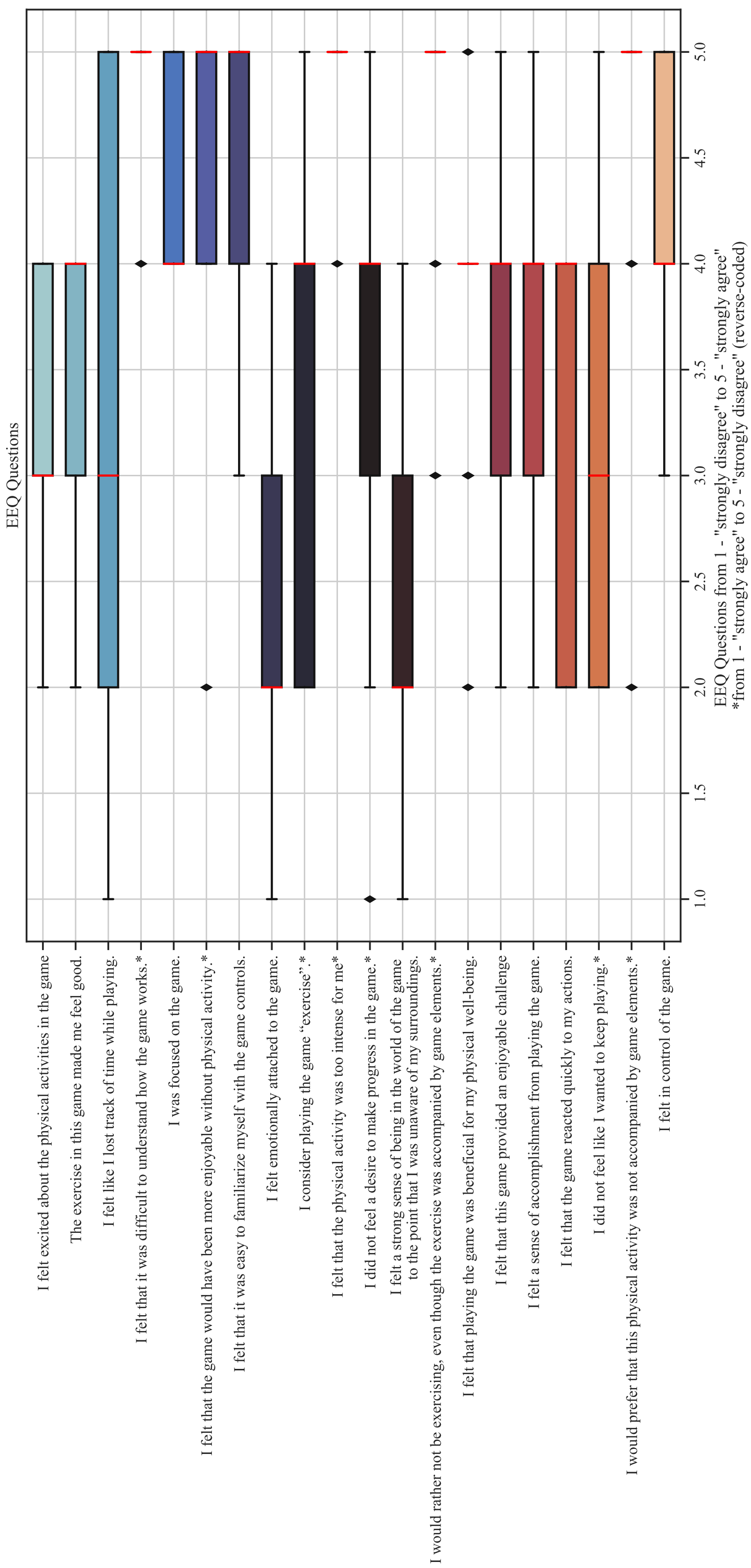
RMSSD Participants Level 1 plots



GEQ Questions boxplot



EEQ Questions boxplot



Think-aloud Statements

participant id	phase	Statement
1	T	Verbindungsaufbau zu lange
1	T	Spurwechsel verwirrend
1	T	Instruktor überlagerte Aktion
1	L1	Kann man beim Wasser weiterspringen?
1	L1	Man kann das Ganze ziemlich bescheißen beim Laufen
1	L1	Da ist eine Lücke im Boden
1	L1	Ist das ein Rundkurs
2	T	It looks like a guy
2	T	I'd say it is confusing that there is grass
2	T	I feel like I am already past the gold (before introductions started)
2	T	Yes! (while finishing)
2	T	I think I am too fast for the game (after the tutorial)
2	T	But Exergame should be slower? So it is fine
2	L1	Sounds are a bit weird, ah it is eating an apple (laughs)
2	L1	(on the questions if sounds should be urned off) No, but 5 apples is a lot
2	L1	Is jump length dependent on the movement?
2	L1	Have to say: I am a little bit bored
2	L1	I just want to jump over the logs and i feel like (the jump) should not be wrong
2	L1	It feels better, when I keep my hands like ths when picking up
2	L1	It was annoying because it felt better when keeping my hands still (picking up)
3	T	Kann ich den jetzt schon nehmen (Apfel)
3	T	Kann ich zurücklaufen (Geld verpasst)
3	T	Wusste nicht, dass man Hindernisse verpassen kann, wenn man zu schnell läuft
3	L1	Zu Sprung: Man muss gar nicht springen?
3	L1	Rechts runter geht nicht
3	L1	Gut, dass man, wenn man keine Ahnung hat was zu tun ist, erst probieren kann und dann nochmal das Tutorial angezeigt wird
4	T	So ein bisschen wie Temple Run
4	T	Geht das auf Zeit oder nur Eingewöhnung
4	T	Hast du das selbst programmiert?
4	T	Bis jetzt sind Interaktionen gut
4	T	Wurde alles erkannt außer greifen rechts unten
4	T	Ob man schneller oder langsamer gehen kann, habe ich nicht rausgefunden
4	L1	Eigentlich könnte man die Einweisung da unten noch größer machen (Ruhig stehen bleiben)
4	L1	Kann ich nicht springen (Stamm)
4	L1	Rechts ist schwierig, rechts mag er nicht (rechts unten)
4	L1	Jetzt muss ich springen, oder? (Graben)
4	L1	Wenn man geht werden nur die Beine oder auch die Arme detektiert, es wäre natürlicher wenn auch die Arme
4	L1	Bekommt man eigentlich irgendwelche Punkte, wenn man die Sachen aufhebt
4	L1	Ich bin angestrengt, weil unten rechts nicht funktioniert (nach Borg)
4	L1	Was passiert,wenn ich einfach weiter ins Wasser laufe
4	L1	Alles so wie vorher aber die Hindernisse kommen schneller
4	L1	Also man kennt jetzt alles und fragt sich was noch passiert
4	L1	Also kommt jetzt wirklich schneller hintereinander, aber kognitiv überfordert bin ich jetzt nicht
4	L1	Man kann sich schon vorstellen, dass das für ältere Menschen irgendwann anstrengend wird
4	L1	Letzter Gedanke: Andere Level geplant?
4	L1	Bisschen langweilig zwischendurch
4	L1	Kann man die Geschwindigkeit selbst steuern?
5	T	Ist ja witzig.
5	T	Auf jeden Fall was anderes als die normalen Psychologiestudien
5	T	Warum war das mit dem Springen faslsch (Ist gesprungen)?
5	L1	Ah nocht nicht (Boxen außer Reichweite)
5	L1	Ah ein großer Schritt stimmt, dann springen beim Fluss
5	L1	Joa ist schon witzig, so kann man die älteren Menschen auch an Gaming heranführen
5	L1	Ich finds gut, dass man den Weg noch weiter sieht,a uch wenn man denkt, das geht jetzt noch ewig
5	L1	Und ich das Gefühl, dass man im Kreis geht, kommt mir so vor, aber das ist nicht so oder?
5	L1	Nett gemacht
5	L1	Erinnert mich an's Winnie Puh Game meiner Schwester
6	T	Es hat einmal einen Schritt nach außen gemacht, automatisch
6	L1	War das jetzt zu früh (nach richtig getimter Aktion)
6	L1	Die optische Wahrnehmung ist bisschen zu nah dran, sodass man schon am Objekt steht, aber nicht greifen kann
6	L1	beim Zurückdrehen mach es immer einen Schritt nach rechts
6	L1	Komisch, dass es den linken Fuß nicht erkennt (Beim drüber steigen)
6	L1	Also beim zurückdrehen mach er automatisch immer wieder einen Schritt zu Seite
6	L1	Den Baumstamm erkennt er immer wieder zu spät beim zweiten Schritt dann
6	L1	Oftmals is es ein bisschen verzögert auch dann beim laufen, aber is gut
7	T	Moorhuhn-Cart-Erinnerung
7	T	Oh Gott (Mücke)
7	T	Gras bei Ziel wegamchen
7	L1	Klatscht bevor es blau wird
7	L1	Fails bei Boxen & Runtergreifen
7	L1	Rechte Sensor geht nicht so gut wie der linke
7	L1	Ich habe ein Problem mit dem Boxen
7	L1	Ah, man muss weiter nach unten boxen
8	T	Das ist ja voll cool
8	T	Ist auf jeden Fall eindrucksvoll
8	T	Leicht verständlioch
8	T	Rechts runtergreifen hat nicht gut geklappt
8	L1	Münzen aufheben ist schwierig
8	L1	Es gibt schnellere und langsamere Phase
8	L1	Spurwechsel geht nicht

9	T	Kann ich das greifen (vor es leuchtet) ?
9	T	Nach unten greifen nicht erkannt
9	L1	Also man darf auf der linken Seite nur mit der linken Hand greifen (Apfel)
9	L1	Mir war klar das (hier) was kommt
9	L1	Nach dem Drehen & Runterbeugen => Laufsperre (Vorschlag)
10	T	Schick!
10	T	Kann ich das Schritttempo beeinflussen?
10	L1	Ah Apfel geht noch nicht (zu früh)
10	L1	Kann es sein, dass ich im Kreis laufe?
11	T	Witzig
11	T	Kann ich schneller laufen?
11	T	Bisschen Delayx zwischen Aktion & Ton
11	T	*lacht*
11	T	Ist nice!
11	L1	Klatschen mehr Delay?
11	L1	Wieso checkt er das?
11	L1	Aufheben geht nicht so gut
11	L1	Geldstapel mögen mich nicht
11	L1	Fluss automatisch ohne Aktion (Arme stark geschlänkert)
12	T	Bisher funktioniert alles gut
12	T	Gut verständlich
12	T	Ist der Berg neu?
12	L1	Holzstamm intuitiv auch springen
12	L1	(Vorschlag) Bach breiter und länger
12	L1	Geschwindigkeit passt nicht immer
12	L1	Funktioniert soweit alles wie erwartet
12	L1	Das Gras wiegt sich im Wind
12	L1	(Vorschlag) Fortschrittsanzeige wie weit im Level
12	L1	Monitor tiefer einstellen, sodass Männchen tiefer ist
13	T	Ah so kann ich mich bewegen
13	T	*Lachen*
13	T	Das ist ja mega cool
13	T	Sehr intuitiv
13	T	gGute Erkennung
13	T	Mega kreativ
13	L1	Dann muss ich wohl loslaufen
13	L1	Fluss oh da muss ich drüberspringen
13	L1	Ich glaube, die Bückbewegung geht langsam besser
13	L1	Wie läuft die Bewertung nach Schnelligkeit, Zeit, Score?
13	L1	(Mücke) vor den Viechern habe ich noch ein bisschen Angst
14	T	ist der Pfeil neu?
14	T	Geht das auch wieder weg von alleine (Instruktor)?
14	T	Funktioniert doch gut!
14	L1	Apfel verwirrend, wenn man drunter steht
14	L1	Gut, dass man ganze Route sieht
14	L1	Bei Äpfeln schwierig räumlich einzuordnen
14	L1	Mücke hat einen Schatten, dass ist gut.
14	L1	Es macht keinen Unterschied auf welcher Spur man ist
14	L1	Ist das ein Rundlauf?
15	T	Ist es wichtig, dass ich die Arme mitbewege?
15	T	Kann ich die Bahn wechseln?
15	T	Macht Spass, bisschen wie Wii
15	T	Farbkennzeichnung auf Pods statt Nummern (Vorschlag)
15	L1	Recht unten geht das Gras in den Bildschirm rein
15	L1	Bin ich zu früh (war zu früh)
15	L1	Was ist das denn (Loch im Boden)
16	T	Schön, dass das Grass sich bewegt
16	T	soll ich durchs Gras durchlaufen?
16	T	Kann ich schon hochgreifen (war noch nicht blau
16	T	Die Töne sind sehr gut
16	T	Räumliche Perspektive bei Apfel ist schwierig
16	T	Feedback, wenn Bewegung zu früh ist wöre gut
16	L1	Bin sehr begeistert von dem Gras
16	L1	Ich bin ein großer Fan von den Soundeffecten
16	L1	Die Boxen mag ich, die sind überraschend.
16	L1	ich denke immer ich bin beim Apfel schon vorbei
16	L1	Mücke ist sehr cool
16	L1	Es geht einen Schritt vor, wenn ich mich bücke
16	L1	Man läuft schon ein bisschen weird auf der Stelle
17	T	Unklar mit zurückdrehen
17	T	Geldmünzen links hat nicht geklappt
17	L1	Gut, dass man schon den ganzen Weg sehen kann
17	L1	Apfel wirkt zu nah, wenn man weg steht
17	L1	Mücken sind witzig
17	L1	Geräusche sind eine gute Unterstützung
17	L1	Wird langsam ein bisschen eintönig (4 min)
17	L1	Was das bei Lücke im Weg?
17	L1	Weltende nicht störend
17	L1	Gras ist beruhigend
17	L1	Immer Statistik zu Äpfeln & Münzen & Boxen (Vorschlag)

Answers to Open Questions

Participant id	Open Question 1	Open Question 2	Open Question 3	Open Question 4	Open Question 5
1	5 min	Ich weiß nicht wie die anderen Level ausehen, aber eher kürzer => zu wenig Challenge.	Egal	War nicht zu komplex.	Nach einiger Zeit => 0.
2	5-10 min	If indeed for exercise => longer, for fun 10 min, but maybe multiple times.	I like how it is, beacuse it is very realistic. I was wondering if you might change the sccenery to the season.	No. The standing still was little weid, but when I understood was fine (figured out in the tutorial) and people would figure that out.	10 min every day if I had to, if it was up to me => I would forget.
3	10 min	maximal 20 min	Lieber 2D, bei 3D war wird mir so schnell schwindelig (war auf VR bezogen), 3D passt.	Nö, ich fands eigentlich ziemlich einfach.	Vielleicht 2x pro Woche.
4	7 min	vielleicht so 4 min	3D, wirkt realer zum Interagieren, vielleicht mehr immersiert => anstrengender.	Nicht komplex, fands gut, weil viele Körperregionen abgedeckt, aufheben war nicht cool, Drehung wäre nice gewesen.	3
5	7-8 min	Nicht so abwechslungsreich => nicht mehr als 10 min.	Wahrscheinlich lieber 3D, wenn man steht, fühlt sich auf jeden Fall natürlicher an.	Nein, gar nicht komplex.	Wenn ich merke, dass ich kognitiv abbaue => lieber spielerisch, dann bei medizinischer Notwendigkeit => 3x Woche.
6	4 min	Ich hätte gerne länger gespielt, weil es Spass geamcht hat => 15 min.	3D war einfach ein bisschen realistischer.	Nein, fand ich nicht => Weil Bewegungen / Interaktionen eindeutig sind.	Freiwillig 1-2x / Woche. Angewiesen alle 2 Tage.
7	10 min	15 min	3D => spannender	Nein.	Wenn medizinisch indiziert 3x / Woche, ansonsten 0x / Woche.
8	7 min	Bis 15 min würde gehen.	3D, macht mehr her, man kann sich Bewegungen besser vorstellen.	Nein.	3x / Woche
9	9 min	5 min	3D, geilere Grafik.	Nö.	Freiwillig nicht. Aber lieber Spielformat als nicht-Spielformat.
10	3,5 min	10 min	3D, weil interessanter.	Münzen schlecht aufgehoben.	Prinzipiell regelmäßig, aber mehr Anreize!
11	7 min	15 min, länger nur mit mehr features.	3D, ist lustiger.	Nö.	0x / Woche freiwillig, bei medizinischer Indikation häufiger
12	4 min	Mit anderen features 15 min, so wie es jetzt ist 10.	3D, immersive => Bisschen anspruchsvoller & ansprechender.	Nein.	Wenn medizinisch indiziert 2x/Tag, sonst nicht.
13	5 min	Weitere Level mit der gleichen Zeit => Insg. 30 min.	3D wirkt echt und effekte sind cooler.	Ne. Bewegungen sind schnell gelernt.	Freiwillig 1x / Woche, wenn medizinisch sinnvoll jeden Tag.
14	7 min	Jetztiger Stand 7 min, eventuell länger mit mehr Features.	3D, leichter vorzustellen mit den Bewegungen, 2D wäre vielleicht zu abstrakt.	Komplex nicht, aber zurückdrehen ist etwas unintuitiv.	Alle 2 Tage mit verschiedenen Levels, täglich bei medizinischen Hinweis.
15	9 min	7 min - bisschen Kürzer, danach entweder neues feature oder aufhören.	3D, man kann sich die Bewegungen besser vorstellen & Besser Bewegungen räumlich planen.	Nö.	Freiwillig: 2x / Woche, medizinisch indiziert 1/ Tag.
16	3 min	10 min	3D, hat mehr Features. Das Gefühl dass man was erkunden kann.	Nein. Intuitive Bewegungen	Freiwillig: Alle 3 Wochen, wenn medizinisch indiziert 2-3 Tage.
17	10 min	5 min mit mehr features oder story.	3D, Vorstellung, dass man in dem Spiel besser ist, Weitblick, Grass beruhigend in 3D. Man kann sich mehr drauf einlassen.	Nein.	Freiwillig 1x / Woche. Medizinisch indiziert 3x/Woche.

List of Figures

List of Tables

Bibliography

[Adc19] Manuela Adcock, Mélanie Fankhauser, Jennifer Post, Kai Lutz, Leopold Zizlsperger, Andreas R. Luft, Vânia Guimarães, Alexandra Schättin, and Eling D. de Bruin. Effects of an in-home multicomponent exergame training on physical functions, cognition, and brain volume of older adults: A randomized controlled trial. *Frontiers in medicine*, 6:321, 2019.

[Adc20] Manuela Adcock, Floriana Sonder, Alexandra Schättin, Federico Gennaro, and Eling D. de Bruin. A usability study of a multicomponent video game-based training for older adults. *European review of aging and physical activity : official journal of the European Group for Research into Elderly and Physical Activity*, 17:3, 2020.

[AH12] Cay Anderson-Hanley, Paul J. Arciero, Adam M. Brickman, Joseph P. Nimon, Naoko Okuma, Sarah C. Westen, Molly E. Merz, Brandt D. Pence, Jeffrey A. Woods, Arthur F. Kramer, and Earl A. Zimmerman. Exergaming and older adult cognition: a cluster randomized clinical trial. *American journal of preventive medicine*, 42(2):109–119, 2012.

[AH18] Cay Anderson-Hanley, Jessica Stark, Kathryn M. Wall, Marisa VanBrakle, Makenzie Michel, Molly Maloney, Nicole Barcelos, Kristina Striegnitz, Brian D. Cohen, and Arthur F. Kramer. The interactive physical and cognitive exercise system (ipaces™): effects of a 3-month in-home pilot clinical trial for mild cognitive impairment and caregivers. *Clinical interventions in aging*, 13:1565–1577, 2018.

[Aho19] Katharina Aholt, Bjoern M. Eskofier, Christine F. Martindale, Arne Kuderle, Heiko Gassner, Till Gladow, Javier Rojo, Samanta Villanueva-Mascato, Jochen Klucken, and Maria Teresa Arredondo Waldmeyer. A mobile solution for rhythmic auditory stimulation gait training. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference*, 2019:309–312, 2019.

[Ame13] *Diagnostic and statistical manual of mental disorders: DSM-5*. American Psychiatric Publishing, Washington, DC, 5. ed. edition, 2013.

[Ame21] American Heart Rate Association. Target heart rates chart, 06.05.2021.

[Amj19] Imran Amjad, Hamza Toor, Imran Khan Niazi, Sanna Pervaiz, Mads Jochumsen, Muhammad Shafique, Heidi Haavik, and Touqeer Ahmed. Xbox 360 kinect cognitive games improve slowness, complexity of eeg, and cognitive functions in subjects with mild cognitive impairment: A randomized control trial. *Games for health journal*, 8(2):144–152, 2019.

[B. 04] B. Winblad, K. Palmer, M. Kivipelto, V. Jelic, L. Fratiglioni, L.–O. Wahlund, A. Nordberg, L. Bäckman, M. Albert, O. Almkvist, H. Arai, H. Basun, K. Blennow, M. De Leon, C. DeCarli, T. Erkinjuntti, E. Giacobini, C. Graff, J. Hardy, C. Jack, A. Jorm, K. Ritchie, C. Van Duijn, P. Visser, and R.C. Petersen. Mild cognitive impairment – beyond controversies, towards a consensus: report of the international working group on mild cognitive impairment. *Journal of internal medicine*, 256(3):240–246, 2004.

[Bam15] Panagiotis D. Bamidis, Patrick Fissler, Sokratis G. Papageorgiou, Vasiliki Zilidou, Evdokimos I. Konstantinidis, Antonis S. Billis, Evangelia Romanopoulou, Maria Karagianni, Ion Beratis, Angeliki Tsapanou, Georgia Tsilikopoulou, Eirini Grigoriadou, Aristea Ladas, Athina Kyrillidou, Anthoula Tsolaki, Christos Frantzidis, Efstathios Sidiropoulos, Anastasios Siountas, Stavroula Matsi, John Papatriantafyllou, Eleni Margioti, Aspasia Nika, Winfried Schlee, Thomas Elbert, Magda Tsolaki, Ana B. Vivas, and Iris-Tatjana Kolassa. Gains in cognition through combined cognitive and physical training: the role of training dosage and severity of neurocognitive disorder. *Frontiers in aging neuroscience*, 7:152, 2015.

[Ban08] Aaron Bangor, Philip T. Kortum, and James T. Miller. An empirical evaluation of the system usability scale. *International Journal of Human-Computer Interaction*, 24(6):574–594, 2008.

[Ben15] Earl T. Benser. Trends in inertial sensors and applications. In Andrei M. Shkel, editor, *2015 IEEE International Symposium on Inertial Sensors and Systems (ISISS 2015)*, pages 1–4, Piscataway, NJ, 2015. IEEE.

[Ben18] Valentin Benzing and Mirko Schmidt. Exergaming for children and adolescents: Strengths, weaknesses, opportunities and threats. *Journal of clinical medicine*, 7(11), 2018.

[Bir15] Marie-Louise Bird, Brodie Clark, Johanna Millar, Sue Whetton, and Stuart Smith. Exposure to ”exergames” increases older adults’ perception of the usefulness of technology for improving health and physical activity: A pilot study. *JMIR serious games*, 3(2):e8, 2015.

[Ble21] Blender Foundation. blender.org - home of the blender project - free and open 3d creation software, 14.05.2021.

[BOR82] GUNNAR A.V. BORG. Psychophysical bases of perceived exertion. *Medicine & Science in Sports & Exercise*, 14(5):377???381, 1982.

[Bro96] J. Brooke. Sus - a quick and dirty usability scale. In Patrick W. Jordan, B. Thomas, Ian Lyall McClelland, and Bernard Weerdmeester, editors, *Usability Evaluation in Industry*. Chapman and Hall/CRC, Boca Raton, 1996.

[Bro13] John Brooke. Sus: a retrospective. *Journal of Usability Studies*, 8:29–40, 2013.

[Bus06] Anja Busse, Matthias C. Angermeyer, and Steffi G. Riedel-Heller. Progression of mild cognitive impairment to dementia: a challenge to current thinking. *The British journal of psychiatry : the journal of mental science*, 189:399–404, 2006.

[But30] Stephen Butterworth. *On the theory of filter amplifiers*. 1930.

[But18] Mary Butler, Ellen McCreedy, Victoria A. Nelson, Priyanka Desai, Edward Ratner, Howard A. Fink, Laura S. Hemmy, J. Riley McCarten, Terry R. Barclay, Michelle Brasure, Heather Davila, and Robert L. Kane. Does cognitive training prevent cognitive decline?: A systematic review. *Annals of internal medicine*, 168(1):63–68, 2018.

[Cer20] Jesus D. Ceron, Christine F. Martindale, Diego M. López, Felix Kluge, and Bjoern M. Eskofier. Indoor trajectory reconstruction of walking, jogging, and running activities based on a foot-mounted inertial pedestrian dead-reckoning system. *Sensors (Basel, Switzerland)*, 20(3), 2020.

[CG_16] CG_Luke. Human base mesh animated. Turbosquid 3D Models, 2016.

[Chu21] Charlene H. Chu, Renée K. Biss, Lara Cooper, Amanda My Linh Quan, and Henrique Matulis. Exergaming platform for older adults residing in long-term care homes: User-centered design, development, and usability study. *JMIR serious games*, 9(1):e22370, 2021.

[Coh88] Jacob Cohen. *Statistical Power Analysis for the Behavioral Sciences*. Elsevier Science, Burlington, 1988.

[Dem20] Jacopo Demurtas, Daniel Schoene, Gabriel Torbahn, Alessandra Marengoni, Giulia Grande, Liye Zou, Mirko Petrovic, Stefania Maggi, Matteo Cesari, Sarah Lamb, Pinar Soysal, Wolfgang Kemmler, Cornel Sieber, Christoph Mueller, Susan D. Shenkin, Lukas Schwingshackl, Lee Smith, and Nicola Veronese. Physical activity and exercise in mild cognitive impairment and dementia: An umbrella review of intervention and observational studies. *Journal of the American Medical Directors Association*, 21(10):1415–1422.e6, 2020.

[Di 07] A. Di Carlo, M. Lamassa, M. Baldereschi, M. Inzitari, E. Scafato, G. Farchi, and D. Inzitari. Cind and mci in the italian elderly: frequency, vascular risk factors, progression to dementia. *Neurology*, 68(22):1909–1916, 2007.

[Dor20] Eva Dorschky, Marlies Nitschke, Christine F. Martindale, Antonie J. van den Bogert, Anne D. Koelewijn, and Bjoern M. Eskofier. Cnn-based estimation of sagittal plane walking and running biomechanics from measured and simulated inertial sensor data. *Frontiers in bioengineering and biotechnology*, 8:604, 2020.

[Dru18] Rafael Rego Drumond, Bruno A. Dorta Marques, Cristina Nader Vasconcelos, and Esteban Clua. Peek - an lstm recurrent network for motion classification from sparse data. In *Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, pages 215–222. SCITEPRESS - Science and Technology Publications, 2018.

[Dub07] Bruno Dubois, Howard H. Feldman, Claudia Jacova, Steven T. DeKosky, Pascale Barberger-Gateau, Jeffrey Cummings, André Delacourte, Douglas Galasko, Serge Gauthier, Gregory Jicha, Kenichi Meguro, John O'Brien, Florence Pasquier, Philippe Robert, Martin Rossor, Steven Salloway, Yaakov Stern, Pieter J. Visser, and Philip Scheltens. Research criteria for the diagnosis of alzheimer's disease: revising the nincds–adrda criteria. *The Lancet Neurology*, 6(8):734–746, 2007.

[Dub10] Bruno Dubois, Howard H. Feldman, Claudia Jacova, Jeffrey L. Cummings, Steven T. DeKosky, Pascale Barberger-Gateau, André Delacourte, Giovanni Frisoni, Nick C. Fox, Douglas Galasko, Serge Gauthier, Harald Hampel, Gregory A. Jicha, Kenichi Meguro, John O'Brien, Florence Pasquier, Philippe Robert, Martin Rossor, Steven Salloway, Marie Sarazin, Leonardo C. de Souza, Yaakov Stern, Pieter J. Visser, and Philip Scheltens. Revising the definition of alzheimer's disease: a new lexicon. *The Lancet Neurology*, 9(11):1118–1127, 2010.

[Dub14] Bruno Dubois, Howard H. Feldman, Claudia Jacova, Harald Hampel, José Luis Molinuevo, Kaj Blennow, Steven T. DeKosky, Serge Gauthier, Dennis Selkoe, Randall Bateman, Stefano Cappa, Sebastian Crutch, Sebastiaan Engelborghs, Giovanni B. Frisoni, Nick C. Fox, Douglas Galasko, Marie-Odile Habert, Gregory A. Jicha, Agneta Nordberg, Florence Pasquier, Gil Rabinovici, Philippe Robert, Christopher Rowe, Stephen Salloway, Marie Sarazin, Stéphane Epelbaum, Leonardo C. de Souza, Bruno Vellas, Pieter J. Visser, Lon Schneider, Yaakov Stern, Philip Scheltens, and Jeffrey L. Cummings. Advancing research diagnostic criteria for alzheimer's disease: the iwg-2 criteria. *The Lancet Neurology*, 13(6):614–629, 2014.

[EG13] Mahmoud El-Gohary, Sean Pearson, James McNames, Martina Mancini, Fay Horak, Sabato Mellone, and Lorenzo Chiari. Continuous monitoring of turning in patients with movement disability. *Sensors (Basel, Switzerland)*, 14(1):356–369, 2013.

[Eli17] ElisaRiva, 2017.

[Fer] Franco Ferraris, Ugo Grimaldi, and Marco Parvis. Procedure for effortless in-field calibration of three-axial rate gyro and accelerometers. *Sensors and Materials*, 1995(Vol. 7):311–330.

[Fit20] Alexander Fitzgerald, Sam Huang, Kyle Sposato, Dongjie Wang, Mark Claypool, and Emmanuel Agu. The exergame enjoyment questionnaire (eeq): An instrument for measuring exergame enjoyment. In Tung Bui, editor, *Proceedings of the 53rd Hawaii International Conference on System Sciences*, Proceedings of the Annual Hawaii International Conference on System Sciences. Hawaii International Conference on System Sciences, 2020.

[GA] Augusto Garcia-Agundez, Ann-Kristin Folkerts, Robert Konrad, Polona Caseman, Stefan Göbel, and Elke Kalbe. Pddancecity: An exergame for patients with idiopathic parkinson's disease and cognitive impairment.

[Gan10] Mary Ganguli, Chung-Chou H. Chang, Beth E. Snitz, Judith A. Saxton, Joni Vanderbilt, and Ching-Wen Lee. Prevalence of mild cognitive impairment by multiple classifications: The monongahela-youghiogheny healthy aging team (myhat) project. *The American journal of geriatric psychiatry : official journal of the American Association for Geriatric Psychiatry*, 18(8):674–683, 2010.

[Gat11] Nicola J. Gates, Perminder S. Sachdev, Maria A. Fiatarone Singh, and Michael Valenzuela. Cognitive and memory training in adults at risk of dementia: a systematic review. *BMC Geriatrics*, 11(1):55, 2011.

[Gil19] Cai Gillis, Fariba Mirzaei, Michele Potashman, M. Arfan Ikram, and Nancy Maserejian. The incidence of mild cognitive impairment: A systematic review and data synthesis. *Alzheimer's & Dementia : Diagnosis, Assessment & Disease Monitoring*, 11:248–256, 2019.

[Gwi19] Piotr Gwiazdowski. Better streaming assets. Unity Asset Store, 2019.

[Hua20] Andrew Hua, Pratik Chaudhari, Nicole Johnson, Joshua Quinton, Bruce Schatz, David Buchner, and Manuel E. Hernandez. Evaluation of machine learning models for classifying upper extremity exercises using inertial measurement unit-based kinematic data. *IEEE journal of biomedical and health informatics*, 24(9):2452–2460, 2020.

[Ira11] Farzin Irani. Visual-spatial ability. In Jeffrey S. Kreutzer, John DeLuca, and Bruce Caplan, editors, *Encyclopedia of Clinical Neuropsychology*, pages 2652–2654. Springer New York, New York, NY, 2011.

[Jen08] Charlene Jennett, Anna L. Cox, Paul Cairns, Samira Dhoparee, Andrew Epps, Tim Tijs, and Alison Walton. Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies*, 66(9):641–661, 2008.

[K. 07] K. Poels, Y.A.W. de Kort, and W.A. IJsselsteijn. *D3.3 : Game Experience Questionnaire: development of a self-report measure to assess the psychological impact of digital games*. Technische Universiteit Eindhoven, 2007.

[Kal60] R. E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1):35–45, 1960.

[Kar17] Esther G. A. Karssemeijer, Justine A. Aaronson, Willem J. Bossers, Tara Smits, Marcel G. M. Olde Rikkert, and Roy P. C. Kessels. Positive effects of combined cognitive and physical exercise training on cognitive function in older adults with mild cognitive impairment or dementia: A meta-analysis. *Ageing Research Reviews*, 40:75–83, 2017.

[Ken91] Deborah Kendzierski and Kenneth J. DeCarlo. Physical activity enjoyment scale: Two validation studies. *Journal of Sport and Exercise Psychology*, 13(1):50–64, 1991.

[Klu17] Felix Kluge, Heiko Gaßner, Julius Hannink, Cristian Pasluosta, Jochen Klucken, and Björn M. Eskofier. Towards mobile gait analysis: Concurrent validity and test-retest reliability of an inertial measurement system for the assessment of spatio-temporal gait parameters. *Sensors (Basel, Switzerland)*, 17(7), 2017.

[Kub] Kubios OY. Kubios hrv standard.

[Kue12] Alexandra M. Kueider, Jeanine M. Parisi, Alden L. Gross, and George W. Rebok. Computerized cognitive training with older adults: a systematic review. *PLOS ONE*, 7(7):e40588, 2012.

[Kui99] J. B. Kuipers, editor. *Quaternions and Rotation Sequences: A Primer with Applications to Orbits, Aerospace and Virtual Reality*. Princeton University Press, Princeton, NJ, 1999.

[Lar02] S. Larrieu, L. Letenneur, J. M. Orgogozo, C. Fabrigoule, H. Amieva, N. Le Carret, P. Barberger-Gateau, and J. F. Dartigues. Incidence and outcome of mild cognitive impairment in a population-based prospective cohort. *Neurology*, 59(10):1594–1599, 2002.

[Lau10] Nicola T. Lautenschlager, Kay Cox, and Alexander F. Kurz. Physical activity and mild cognitive impairment and alzheimer's disease. *Current Neurology and Neuroscience Reports*, 10(5):352–358, 2010.

[Lau19] Nicola T. Lautenschlager, Kay L. Cox, and Kathryn A. Ellis. Physical activity for cognitive health: what advice can we give to older adults with subjective cognitive decline and mild cognitive impairment? *Dialogues in Clinical Neuroscience*, 21(1):61–68, 2019.

[Lui05] H. J. Luinge and P. H. Veltink. Measuring orientation of human body segments using miniature gyroscopes and accelerometers. *Medical & biological engineering & computing*, 43(2):273–282, 2005.

[Lum18] Lumo-Art 3D. Free stylized pbr textures pack. Unity Asset Store, 2018.

[Mad11] Sebastian O. H. Madgwick, Andrew J. L. Harrison, and Andrew Vaidyanathan. Estimation of imu and marg orientation using a gradient descent algorithm. *IEEE ... International Conference on Rehabilitation Robotics : [proceedings]*, 2011:5975346, 2011.

[Mar] Marco Altini. Heart rate variability logger.

[May21] Mayo Clinic. Exercise intensity: How to measure it, 06.05.2021.

[Mig19] Jairo H. Migueles, Alex V. Rowlands, Florian Huber, Séverine Sabia, and Vincent T. van Hees. Ggir: A research community–driven open source r package for generating physical activity and sleep outcomes from multi-day raw accelerometer data. *Journal for the Measurement of Physical Behaviour*, 2(3):188–196, 2019.

[Mit09] A. J. Mitchell and M. Shiri-Feshki. Rate of progression of mild cognitive impairment to dementia–meta-analysis of 41 robust inception cohort studies. *Acta psychiatrica Scandinavica*, 119(4):252–265, 2009.

[Moh10] Mohammad S. Sorower. A literature survey on algorithms for multi-label learning. 2010.

[Mul11] Sean P. Mullen, Erin A. Olson, Siobhan M. Phillips, Amanda N. Szabo, Thomas R. Wójcicki, Emily L. Mailey, Neha P. Gothe, Jason T. Fanning, Arthur F. Kramer, and Edward McAuley. Measuring enjoyment of physical activity in older adults: invariance of the physical activity enjoyment scale (paces) across groups and time. *International Journal of Behavioral Nutrition and Physical Activity*, 8(1):103, 2011.

[Nit19] Nitin J. Sanket. Imu basics, attitude estimation using cf and madgwick, 29.08.2019.

[Nun10] David Nunan, Gavin R. H. Sandercock, and David A. Brodie. A quantitative systematic review of normal values for short-term heart rate variability in healthy adults. *Pacing and clinical electrophysiology : PACE*, 33(11):1407–1417, 2010.

[O'R18] Martin O'Reilly, Brian Caulfield, Tomas Ward, William Johnston, and Cailbhe Doherty. Wearable inertial sensor systems for lower limb exercise detection and evaluation: A systematic review. *Sports medicine (Auckland, N.Z.)*, 48(5):1221–1246, 2018.

[Pet99] R. C. Petersen, G. E. Smith, S. C. Waring, R. J. Ivnik, E. G. Tangalos, and E. Kokmen. Mild cognitive impairment: clinical characterization and outcome. *Archives of neurology*, 56(3):303–308, 1999.

[Pet04] R. C. Petersen. Mild cognitive impairment as a diagnostic entity. *Journal of internal medicine*, 256(3):183–194, 2004.

[Pet10] R. C. Petersen, R. O. Roberts, D. S. Knopman, Y. E. Geda, R. H. Cha, V. S. Pankratz, B. F. Boeve, E. G. Tangalos, R. J. Ivnik, and W. A. Rocca. Prevalence of mild cognitive impairment is higher in men. the mayo clinic study of aging. *Neurology*, 75(10):889–897, 2010.

[Pet11] Ronald C. Petersen. Clinical practice. mild cognitive impairment. *The New England journal of medicine*, 364(23):2227–2234, 2011.

[Pet14] R. C. Petersen, B. Caracciolo, C. Brayne, S. Gauthier, V. Jelic, and L. Fratiglioni. Mild cognitive impairment: a concept in evolution. *Journal of internal medicine*, 275(3):214–228, 2014.

[Pet16] Ronald C. Petersen. Mild cognitive impairment. *Continuum (Minneapolis, Minn.)*, 22(2 Dementia):404–418, 2016.

[por21] Portabiles – customized sensor solutions for wearable and ubiquitous computing, 05.04.2021.

[Raj21] Ashwin Rajkumar, Fabio Vulpi, Satish Reddy Bethi, Preeti Raghavan, and Vikram Kapila. Usability study of wearable inertial sensors for exergames (wise) for movement assessment and exercise. *mHealth*, 7, 2021.

[Rei] Wolfgang Reinhardt, Eva Rügenhagen, and Bernhard Rummel. System usability scale - jetzt auch auf deutsch.

[Ren20] Render Knight. Fantasy skybox free. Unity Asset Store, 2020.

[RO17] Carolina Rico-Olarte, Diego M. López, Santiago Narváez, Charic D. Farinango, and Peter S. Pharow. Haphop-physio: a computer game to support cognitive therapies in children. *Psychology research and behavior management*, 10:209–217, 2017.

[Rob14] Rosebud O. Roberts, David S. Knopman, Michelle M. Mielke, Ruth H. Cha, V. Shane Pankratz, Teresa J. H. Christianson, Yonas E. Geda, Bradley F. Boeve, Robert J. Ivnik, Eric G. Tangalos, Walter A. Rocca, and Ronald C. Petersen. Higher risk of progression to dementia in mild cognitive impairment cases who revert to normal. *Neurology*, 82(4):317–325, 2014.

[Row16] Alex V. Rowlands, Tom Yates, Melanie Davies, Kamlesh Khunti, and Charlotte L. Edwardson. Raw accelerometer data analysis with ggir r-package: Does accelerometer brand matter? *Medicine and science in sports and exercise*, 48(10):1935–1941, 2016.

[Ruea] Christoph Ruegg and Markus Cuda. Mathnet.numerics. NuGet.

[Rueb] Christoph Ruegg, Matthew Kitchin, Thomas Ibel, and Albert-Jan Nijburg. Mathnet.filtering. NuGet.

[Sab14] Angelo Maria Sabatini and Vincenzo Genovese. A sensor fusion method for tracking vertical velocity and height based on inertial and barometric altimeter measurements. *Sensors (Basel, Switzerland)*, 14(8):13324–13347, 2014.

[Sch13] Daniel Schoene, Stephen R. Lord, Kim Delbaere, Connie Severino, Thomas A. Davies, and Stuart T. Smith. A randomized controlled pilot study of home-based step training in older people using videogame technology. *PLOS ONE*, 8(3):e57734, 2013.

[Sha20] Ilc Shatalmic. Bluetooth le for ios, tvos and android. Unity Asset Store, 2020.

[Tso07] Grigorios Tsoumakas and Ioannis Katakis. Multi-label classification. *International Journal of Data Warehousing and Mining*, 3(3):1–13, 2007.

[Uni20] Unity Technologies. Standard assets (for unity 2018.4). Unity Asser Store, 2020.

[van94] Maarten W. van Someren, Yvonne F. Barnard, and Jacobijn A. C. Sandberg. *The think aloud method: A practical guide to modelling cognitive processes*. Academic Press, London, 1994.

[van14] Vincent T. van Hees, Zhou Fang, Joss Langford, Felix Assah, Anwar Mohammad, da Silva, Inacio C. M., Michael I. Trenell, Tom White, Nicholas J. Wareham, and Søren Brage. Autocalibration of accelerometer data for free-living physical activity assessment using local gravity and temperature: an evaluation on four continents. *Journal of Applied Physiology*, 117(7):738–744, 2014.

[W.A13] W.A. IJsselsteijn, Y.A.W. de Kort, and K. Poels. *The Game Experience Questionnaire*. Technische Universiteit Eindhoven, 2013.

[Wol20] Bettina Wollesen, Alicia Wildbredt, Kimberley S. van Schooten, Mei Ling Lim, and Kim Delbaere. The effects of cognitive-motor training interventions on executive functions in older people: a systematic review and meta-analysis. *European review of aging and physical activity : official journal of the European Group for Research into Elderly and Physical Activity*, 17:9, 2020.

[Wor19] World Health Organization. 10 facts on dementia, 1/09/2019.

[x-i21] x-io Technologies. Open source ahrs with x-imu – x-io technologies, 08.05.2021.

[Zha20] Yinan Zhao, Hui Feng, Xinyin Wu, Yan Du, Xiufen Yang, Mingyue Hu, Hongting Ning, Lulu Liao, Huijing Chen, and Yishan Zhao. Effectiveness of exergaming in improving cognitive and physical function in people with mild cognitive impairment or dementia: Systematic review. *JMIR serious games*, 8(2):e16841, 2020.