Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Computer Science

Nikolai Åkerholt

Evaluating a remote data collection platform: A practical user study on exercise trackers

Master's thesis in Applied Computer Science
Supervisor: Mariusz Nowostawski
June 2021

NTNU
Kunnskap for en bedre verden

Nikolai Åkerholt

Evaluating a remote data collection platform: A practical user study on exercise trackers

Master's thesis in Applied Computer Science
Supervisor: Mariusz Nowostawski
June 2021

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Computer Science

NTNU
Norwegian University of
Science and Technology

Preface

This master thesis was written as part of a Master's degree in Applied Computer Science at NTNU Gjøvik. My work in this thesis is split into two parts. The first part presents a practical example of a remote data collection platform, which was found even more valuable to enable the possibility to run Virtual Reality (VR) exergame experiments despite the Covid-19 pandemic. The second part is an investigation on the gamification provided by exercise trackers, and their effect on people's engagement with physical activity.

I chose the topic of exercise trackers as I personally have a large interest in how gamification can be used to promote serious purposes, such as making our society make more healthy life choices. Further, due to the physical restrictions caused by Covid-19, I saw this as an opportunity to explore technologies to enable remote user studies.

The thesis is aimed at a general computer science audience, both academics, and practitioners. However, I think the thesis is relevant for a wider audience, as anyone interested in remote user study systems or gamification to promote exercise can get something valuable out of this thesis.

01-06-2021

Acknowledgements

This thesis would not have been possible to pursue without the considerable support and help I have received from others.

I would like to thank my supervisor, Mariusz Nowostawski, for his excellent counselling and academic insights throughout the master thesis project.

I would like to thank my girlfriend, Ida Sofie Holmen, for her love and support through my process of writing this thesis.

I want to thank my family and friends for their continuous support and interesting discussions related to my thesis topic.

I would like to thank Morten Omholt-Jensen, for our initial ideation and discussions related to remote rehabilitation systems which lead to the topic I chose to pursue.

I want to express my gratitude to all the participants for their concerted efforts and feedback in the two-week user study. I also want to thank everyone who took their time to respond to my online research questionnaire.

This research was partially funded from personal funding, and from NTNU. I want to thank NTNU for their provided funding; without it, the research experiments would have been more difficult to arrange.

Lastly, I want to give a big thank you to all the professors, students and other people I came across during my 5 years here at NTNU Gjøvik. You have made my studies here an amazing and memorable experience.

Nikolai Åkerholt

Abstract

This thesis investigates two topics: The first topic is remote data collection for user studies, a topic deemed important recently due to the physical implications caused by the Covid-19 pandemic. The second topic is gamification techniques for exercise motivation, a popular subject in medical research, but also very important for our sedentary society. Physical inactivity is a worldwide issue, and attention on this topic can be crucial to improve our general well-being and life quality.

Throughout the thesis, I have investigated the topics through two focus areas:

Focus A: Technologies to enable remote user studies

Focus B: Gamification techniques for exercise motivation

I investigated Focus A through a literature review and practical implementation of a remote user study system called *ExerIsland*. I created the system to enable remote user studies on an exergame platform in VR. My technical solution provided a VR Exergame platform for participants to interact with minigames, while data about their activities are being logged. Participants and researchers have access to an online web dashboard for displaying this data. In the end, I tested my ExerIsland system with a practical two-week remote user study with 7 participants. I arrive at the conclusion that my system was successful in enabling remote user studies based on findings and personal experiences. However, the presented solution has some flaws in terms of security and facilitation.

Focus B was also investigated through a literature review and through results of the two-week user study with *ExerIsland*. In addition, I distributed an online questionnaire related to this topic, which received 94 responses. My findings indicate that people using exercise trackers deem the exercise activity itself as *fun* more often. Additionally, there is a connection between exercise tracker usage and average workout hours per week. I further discovered that people have varied opinions on how they like to inspect data from exercise logging. Personal statistics to compare with oneself was one of the popular ways. The findings from my study are not enough to display any causal relations directly implied by exercise trackers. My results serve only as an indication of the potential benefits given by the simple gamification provided by the tested exercise tracking.

Contents

List of Figures

List of Tables

Listings

Acronyms

CRUD Create, Read, Update and Delete. 38

CSV Comma-separated values. 76

DDoS Distributed Denial-of-Service. 39

IoT Internet-of-Things. 22

MVP Minimum-Viable-Product. 32

NCD non-communicable disease. 1

QoL Quality-of-Life. 62

SDK Software Development Kit. 33

SDT self-determination theory. 2

SLOC Source Lines of Code. 31

UI User Interface. 58

VR Virtual Reality. i, iii, 2, 10, 13, 22, 103

WHO World Health Organization. 1

Terms

Brute-force attack is a security attack where the user continuously tries many combinations of credentials at the same time. The malicious users hope that one of the combinations will get a hit, providing access to something they should not have. 104

Coroutine is a type of function that can pause its execution and give control back to Unity. At a later point in runtime, it will continue where it left off on the following frame. This is useful for running asynchronous tasks without multithreading [1]. 64

Dart is a client-optimized programming language for apps on multiple platforms. It is developed by Google and is used to build mobile, desktop, server, and web applications. 33, 58

Exergame is a type of serious game with a main focus on promoting or inducing physical exercise. 2, 22, 25, 28, 45

Flow is a state of optimal experience arising from intense involvement in an activity that is enjoyable. Flow arises when one's skills are fully utilized yet equal to the demands of the task, and intrinsic motivation is at a peak (proposed by Mihaly Csikszentmihalyi in 1990) [2]. 47, 52

Flutter is an open-source UI software development kit created by Google. It is designed to create natively compiled applications for mobile, web, and desktop from a single codebase. 33, 58

Gamification is the use of game design elements in non-game contexts. 1

Golang is an open-source programming language developed by Google. Often used in back-end architectures and systems. 33

Immersion (in context of Virtual Reality) is a person's perception of physical presence in a non-physical world. 50

Kinect Camera and motion sensor from Microsoft, no longer in production. 22

Prefab (in context of Unity) is a prototype object template that can be saved and instantiated in game scenes. 42

Serious Game is a game with a specific primary purpose other than pure entertainment. 1, 9, 50

Spoofing is a security attack where a user pretends to be someone else. 104

Telerehabilitation is a form of rehabilitation where the patients are performing their exercises remotely, and communicating with their providers remotely. 2, 5

Unity 3D Game Engine is a game engine often used to develop both 3D and 2D games. Uses C# and .NET. 32, 45

1 Introduction

1.1 Topic

Health and wellness are both important in our daily lives. Everyone cares about and wants to look after their health to different degrees, as it is vital for having a long and happy life. Having good health does not only restrict itself to one thing we need to look out for. It is a series of things, that in combination makes up both our physical and mental health.

> *Health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity.*
>
> - *World Health Organization (WHO)* [3]

To sustain good health, there are basic needs that have to be fulfilled and actions we need to make as individuals. Some of these things include; eating food that is good for us, sleeping enough to feel well-rested, and performing some physical activity for our wellness both physically and mentally. Physical activity can be defined as "*any bodily movement produced by skeletal muscles that results in energy expenditure*" [4]. Although the need for physical exertion is not as basic and obvious to us as eating and resting, as those basic needs are felt very directly in our bodies. If your body needs energy from food, you can feel hunger. If your body needs to rest, you feel tiredness or exhaustion. However, when it comes to physical inactivity, our bodies do not tell us in the same obvious way that we need to do something about it. You may feel an urge to exercise or do some physical activity if you have been sitting still for too long, but your body is telling this in a much more vague way. In combination with the fact that our technological leaps in society have made us less dependent on physical activity, these could be substantial factors as to why physical inactivity is the fourth leading cause of death worldwide [5].

Physical inactivity can lead to non-communicable diseases (NCDs), which is a major burden worldwide. [6] NCDs are also known as chronic diseases, and they tend to be of long duration and are a result of a combination of genetic, physiological, environmental and behavioural factors [7]. Although efforts are being done globally to reduce the general risk of NCDs in society, there has been less attention given to the importance of an active lifestyle for disease prevention [6].

In the context of applied computer science, many technological interventions are being developed and researched to help us achieve better health. In the case of physical health and exercise, it also applies there. To achieve long-term engagement in physical activity, it could be very important to reinforce people's intrinsic motivation, through interest and enjoyment [8]. One potential way to increase interest and enjoyment in activities is through the use of gamification or serious games. Gamifying non-gaming activities have become increasingly popular to increase engagement and enjoyment in users [9].

Gamification is the use of game design elements in non-game contexts.

- Sebastian Deterding et al. [10]

Gamification can be applied simple or in an advanced fashion. In fact, many of us apply gamification in daily life activities without even thinking about it. For example; creating a to-do list is comparable to receiving a set of missions to complete in a game. Ticking off one item in that list gives us a very simple, yet rewarding sensation when done. The fact that we tend to gamify many of our day-to-day activities, describes our need for some extra sensation of reward, or motivation when it comes to doing certain activities. That type of gamification can be transferable to many different activities in life, including *exercise*. I wanted to explore the topic of using gamification by the means of exercise tracking to motivate people to engage in physical activity.

Moreover, I wanted to research and develop a system to conduct remote data collection in a user study with an exergame application. The physical constraints caused by Covid-19 further emphasized the importance of this topic. The tool would be very useful, as any type of on-site physical studies currently can be hard to pursue due to the physical implications created by the pandemic.

In other words, the thesis aims to investigate *how to create a data collection platform for remote user studies*. In addition, the thesis will evaluate how gamification elements can be applied to increase our engagement with physical activity.

1.2 Keywords

Remote user study, remote systems, remote data collection, gamification, exercise tracking, exergames, VR, physical activity, exercise motivation.

1.3 Motivations

On a personal level, I am interested in the topic of human motivations and what factors drive us towards making decisions we make. It is interesting to try to understand how external factors, such as extrinsic rewards or interaction from other people, can potentially cause an initial spark of motivation to take life-changing actions. In many cases, this extrinsic motivation may falter over time and the initial spark was all there ever was. However, there are times when this initial spark transfers to intrinsic motivation, which in turn, could be much more effective at keeping people engaged in something in the long-term [11]. This is where it can be interesting to understand how self-determination theory (SDT) and people's different personality traits can matter when it comes to getting the intrinsic motivation to do physical exercise.

Personally, I have also experienced these motivating mechanisms that have led to increased physical activity. However, I also acknowledge that my motivation for physical activity is inconsistent and does change over time. In combination with the importance in current society and research, I deemed the topic as very interesting to further investigate.

The initial spark of motivation to this project were two things; the large wave of elderly retiring in the upcoming years and the Covid-19 pandemic. Together with a co-student, I previously came up with the idea of a telerehabilitation system. It was meant for people in care centres that needed rehabilitation but were restricted to meeting physically. Either due to the lack of availability in the rehabilitation centres, or more recently, because of the pandemic.

When I researched that topic, I found it interesting to understand what elements cause positive changes in people's engagement with physical activity. Further encouraged by the fact that it is a large societal issue that we are too physically inactive in general. [5]

I realized that the relation between the *health-care professional and the patient* in a telerehabilitation system is very similar to a *research facilitator and study participant* in a remote user study setting. As the relevance for a *remote study system* were even greater due to the pandemic, it further motivated me to pursue this topic.

1.4 Focus Areas

The goal of this thesis is to investigate how a system can be created to enable remote user studies, and how gamification can increase our physical activity levels. To facilitate reaching these goals, two *focus areas* were defined:

Focus A:

Technologies to enable remote user studies

Focus B:

Gamification techniques for exercise motivation

1.5 Research Questions

To further specify and narrow the focus of the thesis, there were three research questions defined. Research question 1 is directly relevant to focus A, while research question 2 and 3 are applicable to focus B.

Research Question 1:
How to develop an exergame platform that enables the ability to run remote user studies?

Research Question 2:
How can exercise trackers motivate people to engage more in physical activity?

Research Question 3:
How should data from games be presented to engage users more in activities?

1.6 Contributions

The main contributions of this thesis can be divided into three main parts:

A **literature study** (chapter 2) was conducted to investigate the current state-of-the-art and research within the field of both *remote user studies*, as well as investigating *gamification* and its effect on exercise and general task motivation.

Secondly, an **online research questionnaire** was distributed to collect data about peoples' tendencies and subjective thoughts on exercise trackers and how they affect their physical activities.

Lastly, a **two-week user study** was conducted with a practical example of a developed exergame platform. The platform contains minigames to play in VR and an online dashboard on web to track game statistics.

This practical prototype used to conduct the two-week study experiment is the main developed artifact of the thesis. Design and development details of the platform are explained in chapter 3. Finally, the experiment methodology utilized with the platform is explained in chapter 4.

1.7 Thesis Structure

The chapters of this thesis do not follow a strictly traditional IMRAD structure. However, I would argue that the contents of the thesis are following an IMRAD-like layout. The thesis begins with an introduction (chapter 1), then follows up with a further explanation of the background and existing literature in chapter 2. After that, the document follows a structure which is more accurately describing the order of how contributions were made in the study. This is done to tell a more cohesive story, while hopefully also providing a better reading experience. Below is a list of the chapters with short descriptions to give the reader an easy overview of the thesis.

Chapter 1: Introduction gives an overview of the contents and goals of the thesis.

Chapter 2: Background describes the literature study conducted to get an overview of the current state-of-the-art within the research area. This chapter contains method, results and a summary of the findings at the end.

Chapter 3: ExerIsland: An Exergame Platform for Remote User Studies describes the developed exergame platform. Detailing the system architecture, design, and technical implementation details that enabled the possibility to run remote study experiments.

Chapter 4: Experiment Methodology explains the experiment methodology that was utilized both in the *two-week user study* as well as the *online research questionnaire*.

Chapter 5: Experiment Results details the different findings and specific results of the *two-week user study* and the *online research questionnaire*.

Chapter 6: Discussion describes my personal interpretations on how results are relevant in regards to the defined research questions.

Chapter 7: Summary & Conclusions covers the final conclusions of the thesis.

2 Background

To get a better understanding of the current state-of-the-art for the present research topics, I wanted to conduct a literature review in respect to this area. More specifically, the goals of this literature study were divided into two parts.

The first part was about investigating how to create a system that enables *remote user studies*. The second part was about exploring gamifications being researched related to increasing peoples motivation or performance in physical activity.

This literature study would help identify specific gaps in current state-of-the-art, and potentially find interesting knowledge and lessons learned from other research projects. In turn, this would be very valuable for me to utilize in the development of the ExerIsland platform. Moreover, to shape the methodology for conducting experiments with the platform.

Note: parts of the content and results from this chapter is based on my work conducted in IMT4134 Specialization in Software Engineering. It has been modified and extended for the context of this thesis.

2.1 Focus

In both the medical domain and general health and wellness, several research projects are being conducted in terms of utilizing new software systems and technology as a way to improve rehabilitation or exercise quality. In Norway specifically, Sunnaas Hospital [12] is one of the larger entities in the rehabilitation space when it comes to applying new technology in their rehabilitation programs. Their patients often have severe issues that require advanced and prolonged treatment to recover. They are using games and also VR as part of their rehabilitation programs. Moreover, they have created a website called *Spilldegbedre.no* [13] which can recommend anyone to off-the-shelves entertainment games for home rehabilitation. There also exist some commercialized products using VR and gamification as therapy. For example, *Psious* [14] is a company that creates games and experiences to help people overcome anxieties and phobias.

In regards to remote systems, there is a lot of research being done in this space currently, especially for telerehabilitation systems in the medical domain. A review of the state-of-the-art in telerehabilitation was conducted by A. Peretti et al. [15] and they describe that there exist several good examples of applications where telerehabilitation can be applied, but that there are many gaps to be filled. They emphasize that it is important to include the users when creating requirements for such software. Having said that, the literature seems to be limited in regards to how remote *user studies* can be conducted. This is why I deem it important to further investigate that subject area with this literature review.

In regards to looking at software technology or gamifications applied to promote exercise, there was a systematic literature review published in 2019 by J. Koivisto and J. Hamari [16]. In their study, they ended with a total of 16 articles in their synthesis, and they found some common trends with current research. Their findings indicate that the most recurring gamification intervention to induce physical activity was reward systems. Such as points, leaderboards, goals, and progress visualization tools. Additionally, they state that the general outcome of their reviewed literature tends to be positively oriented, meaning that gamification elements seem to positively affect physical activity. However, they also state that more research with controlled study settings and using more objective validated ways of measuring outcomes in this area would be beneficial [16]. Many articles seem to go in the direction of self-reporting personal physical activity, and this is used as an indicator for the actual outcome. This is cost-efficient and simple, but it is not an objective way of measuring. The findings of J. Koivisto and J. Hamari [16] will be used for comparison in this systematic literature review, as it is in a similar topic area. In a way, it can give common or contradicting perspectives, meaning that it will serve as a background to either confirm or disconfirm potential findings in this study.

Focus Areas

When the literature study was conducted, it was done within the two focus areas defined for this thesis:

Focus A:

Technologies to enable remote user studies

Focus B:

Gamification techniques for exercise motivation

The idea was that the results from the literature study would provide new knowledge to make better decisions for the artifacts developed and utilized in this thesis. Focus A would give relevant insights into how to develop the platform in a good way to enable remote study experiments, directly relevant to research question 1 of the thesis: "*How to develop an exergame platform that enables the ability to run remote user studies?*"

Focus B, on the other hand, would give more valuable information that was more related to the practical experiment conducted, which was investigating the motivational effects of gamification. Meaning, that it was more relevant for research question 2: "*How can exercise trackers motivate people to engage more in physical activity?*" and research question 3: "*How should data from games be presented to engage users more in activities?*" of the thesis.

In turn, this literature study would not only have an impact on the outcome of the experiments but also on how useful the remote study platform would become in solving its goal.

2.2 Method

This section describes the planning phase I had while conducting the systematic literature review. It includes motivations for the focus area, search strategy, inclusion- and exclusion criteria, and also paper classification methodology.

The methodology used to conduct the Focus B literature study follows a quite structured systematic literature review approach. Focus A, on the other hand, follows a somewhat less strict literature review approach. The reason for this is that there were a lot less literature directly relevant to how remote user studies could be conducted. This made it beneficial to follow a less strict literature review methodology for this focus area. However, the methodologies are similar and are documented below.

2.2.1 Focus A: Technologies to enable remote user studies

Search strategy

First of all, when coming up with a search strategy, a set of relevant keywords was defined. This way, I had some options to test out different combinations to see what gave the best and relevant results to the focus area defined in section 2.1.

For the literature study regarding remote technologies, this was the set of keywords that were used in the search for literature: *Remote study platform, user study, software development, technology, researcher, participant, data-collection, remote user study, remote monitoring technology, VR.*

These keywords were used in different combinations, and as there were not that much directly relevant literature to this topic area, there was some variance in when the literature was published. Included literature range between 2007-2021 in terms of publishing date. The three main databases that were used for finding literature was SpringerLink[1], IEEE Xplore[2] and Google Scholar[3]. Although Google Scholar provided a high number of results in the queries, some of the first pages did result in some relevant findings. Therefore, the first few pages of results were further investigated and included. The full queries, literature evaluated and included can be seen in table 1 below.

Search string	Filters	Database	Results	Articles evaluated	Articles included
(User Study) AND (remote study platform) AND (software development) AND technology AND researcher AND participant AND data-collection	Year: 2019-2021 Discipline: Computer Science	SpringerLink	248	10	4
(User Study) AND (remote study platform) AND (software development) AND technology	Year: 2014-2021	IEEE Xplore	26	3	1
Remote user study	None (only looked through first 3 pages)	Google Scholar	3.990k+	4	2
(Remote monitoring technology) AND VR	Year: 2014-2021 (only looked through first 3 pages)	Google Scholar	20k+	4	3

Table 1: Focus A: Search queries, number of results and included articles

[1] link.springer.com
[2] ieeexplore.ieee.org
[3] scholar.google.com

Inclusion and exclusion criteria

This literature study initially started with quite strict inclusion and exclusion criteria. The aim was to only find papers that had created systems for remote user study experiments. However, when conducting the literature search, I found out that this topical area seemed to be quite a niche topic, and there was not enough literature out there to justify this.

When searching for *remote experiments*, most results yielded was in regards to in-lab experiments conducted for physicists and similar lab environments. In this literature review, on the other hand, the topic area was to investigate specifically remote user studies and not general lab experiments enabled by remote access.

Due to the limited amount of literature, the inclusion- and exclusion criteria ended up being quite open to any literature that was relevant to anything with remote study experiments or remote tracking systems:

Inclusion Criteria

IC1: Has objective of investigating technology or methodology to enable remote user study experiments.

IC2: Has described objective of investigating remote tracking systems.

Exclusion Criteria

EC1: The study is not written in English.

EC2: Study does not address the topic of either remote user studies or remote systems.

EC3: Full-text of paper is not accessible.

Paper classifications

Some simple classifications was made in regards to Focus A (see table 2). This way, I could more easily categorize and synthesize data together in a spreadsheet. Making it easier to look up and collect some general insights on the literature out there. Additionally, in my subjective opinion; with these classifications, it is easier to quickly identify specific findings of the literature reviewed.

Property	Categories
Study-data	Qualitative, Quantitative, Mixed, Literature Review
Research approach	Survey, Experiment, Case-study, R&D, Literature Study
Hardware Technologies	VR, Computer, Camera tracking, Videogaming console, Mobile, Wearable-technology, Mobile sensors
Research topic(s)	Remote Monitoring, Remote User Studies, IoT, Logging, VR, Machine Learning, Rehabilitation
Software Technologies	(No strict classifications, but notes about programming language, paradigms, software architectures, development frameworks, etc.)
Remote context	Rehabilitation, Experiment, Medical, Industry, Observation
Display of data	Session graphs, Metric over time graphs, Graphical/Video, data tables
Metrics recorded	(No strict classifications, but notes on what data their remote system collected. Such as position, bodily movement, score in games, subjective participant data, etc)

Table 2: Focus A: Paper Classifications - Properties and categories

Execution

In the execution of this literature study, I applied my search methodology and went through papers to investigate whether they were relevant to further review. Conducting the initial queries resulted in too many results to explore every single one. However, this was not seen as possible or intended either as the focus area seemed to have a limited amount of directly relevant research.

A total of *21* papers were further read after initially evaluating the title and abstract. After inclusion and exclusion criteria was applied, I was left with ***10*** [17–26] papers which were found relevant for Focus A of this literature review. The spreadsheet containing the final papers can be seen in appendix A.

2.2.2 Focus B: Gamification techniques for exercise motivation

Search strategy

In Focus B, there were much more relevant results as there were more studies conducted within this topic area. To narrow the search down to get more appropriate results, a larger set of keywords were defined here compared to the one used for Focus A.

These are the keywords that were defined: *Software engineering, techniques, technology, gamification, serious games, exergames, game, application, exercise, training, workout, rehabilitation, performance, physical health, intervention, effects, long-term, motivation, motivational effects.*

These keywords were used in a few different contexts to create relevant, narrowing search strings on the different databases used. Additionally, only papers from 2014-

2021 were queried for, as those were found to be most relevant (described below in section 2.2.2). The two databases that were used to query for relevant literature were IEEE Xplore and SpringerLink. Other databases, such as Google Scholar and ACM Digital Library[4] were also tested out in querying for papers. However, these yielded way too many results with simple queries[5](∼24,000 on Google Scholar and ∼17,000 on ACM Digital Library).

In addition to these databases yielding too many results, I was also more familiar with querying with IEEE Xplore and SpringerLink and did not want to create too much overhead work in the solo literature review. In the end, the final decision was made to stick with these two databases to query for papers. Table 3 displays the full query string and filters which were used.

Search string	Filters	Database	Results	Articles evaluated	Articles included
(Software engineering OR techniques OR technology) AND (gamification OR serious games OR exergames OR game OR application) AND (exercise OR training OR workout OR rehabilitation) AND (physical health OR intervention OR effects OR long-term OR performance) AND (motivation OR motivating OR engaging OR engagement)	Year: 2014-2021	IEEE Xplore	464	38	12
(Software engineering OR techniques OR technology) AND (gamification OR serious games OR exergames OR game OR application) AND (exercise OR training OR workout OR rehabilitation) AND (physical health OR intervention OR effects OR long-term OR performance) AND (motivation OR motivating OR engaging OR engagement) AND (participant OR participants OR subject OR subjects)	Year: 2014-2021	IEEE Xplore	92	17	7
(Software engineering) AND (exercise OR rehabilitation) AND (physical health) AND (intervention OR effects OR long-term OR performance) AND (motivation OR engagement) AND participants AND experiment	Year: 2014-2021 Discipline: Computer Science	SpringerLink	487	21	9

Table 3: Focus B: Search queries, number of results and included articles

Inclusion and exclusion criteria

In this study, only papers from 2014-2021 were included. This is because the different technologies applied for exercise motivation are massively changing in terms of how advanced they are. For example; if a paper published before 2010 were looking at VR as an intervention for exercise, it is not necessarily comparable to a more recent study, because VR technology has changed so much in recent years. It is also usually better to look at more recent research when conducting these reviews, to get a more proper view of the current state-of-the-art.

Additionally, when conducting the study there were defined a few inclusion- and exclusion criteria for deciding whether the articles would be included or not in the final study:

Inclusion Criteria

IC1: Has primary objective of looking at interventions to affect motivation for physical exercise.

IC2: Has described secondary objective of looking at physical exercise motivation.

IC3: Has described primary objective of looking at motivational elements to perform a specific activity (other than physical exercise).

IC4: Has conducted practical study, testing their intervention in practice.

[4] dl.acm.org

[5] Example query used: Software engineering AND (gamification OR game) AND (exercise OR training OR workout OR rehabilitation) AND physical health

Exclusion Criteria

EC1: The study is not written in English.

EC2: Study does not address the topic of either exercise or motivation.

EC3: Full-text of paper is not accessible.

EC4: Summaries of abstracts or larger reviews that does not describe a specific practical study.

Paper classifications

When it came to classifications of papers, I created a few categories where it would be relevant to concretize information into specific branches. Later on, this information could be used to potentially give some valuable information in regards to focus B (see section 2.1). The properties and different categories can be seen in table 4 below.

Property	Categories
Study-data	Qualitative, Quantitative, Mixed
Research approach	Survey, Experiment, Case-study
Technology used	VR, AR, Computer, Camera body tracking, Gaming console, Mobile application, Tablet, Wearable-technology, Feet trackpad, Physical instrument
Research topic	Exercise motivation, Task performance, User experience, Usability, Task motivation, Physical activity
Intervention	Competition, Self-competition, Reward-system, Performance feedback, Exergame, Game, VR Immersion, Autonomy, Adjusting difficulty, Socialization, Music
Motivation results	Increased, Increased then fading, Mixed, None, Declined
Long-term motivation results	12 months or longer, 6 months or longer, 3 months or longer, No.
Other results	Increased physical performance, Increased cognitive performance, Increased social interaction, Increased task performance, Increased well-being

Table 4: Focus B: Paper Classifications - Properties and categories

Execution

In this phase, I applied my search methodology. The number of article results from all queries was *1043* in total. From all of these, only *76* were chosen to be more thoroughly read through. There were rather many articles that were discarded from these queries in total, and there were two reasons for this:

The first reason was that the first items resulting from the search engines usually were the most relevant, and accurately going into the correct topic area. However, when going through further results from the search engines, it became clear that the results became less relevant. This is likely due to the algorithm utilized by these search engines, providing the results that fit the queries the best first, and then less precise results later. Due to this, I could recognize that many papers were not suiting the research topic for this literature study, solely based on the *title*.

Secondly, this review was conducted as a solo review. It created some limitations in the workload I could persist for a solo review. As the major part of relevant papers came from the first parts of the results from the search engines, I had to make a hard limit on how many papers to go through in order to not create too much workload. This is further described in section 2.2.3 below.

After more thoroughly reading and investigating into these 76 papers, only ***28*** papers total [27–54] remained to be included for the systematic literature review after applying inclusion and exclusion criteria. Spreadsheet containing the final papers can be seen in appendix A.

2.2.3 Threats to validity

As depicted in tables 1 and 3, there were many article-results from the search queries conducted in total. As this literature review was only conducted by myself, there were some rather large limitations in terms of workload. I could not be able to look through all of the titles and abstracts of the article results. In afterthought, I think the initial search queries should have been more narrow. Especially in regards to Focus B where there was a more systematic literature review approach, and there were also much more relevant research articles within the topic area. For Focus A it is also a limitation, but it was almost required to not follow a too strict methodology for that focus area to not have a too limited amount of literature.

Additionally, I think that it would have been useful to have second opinions from a co-student in this project. These issues are further described in section 2.4. However, I think that the articles included are relevant to provide useful information related to shaping the study platform. Moreover, to provide useful indications for the research questions. They provide an indication of current research within the area of remote user studies, and software engineering techniques for exercise motivation.

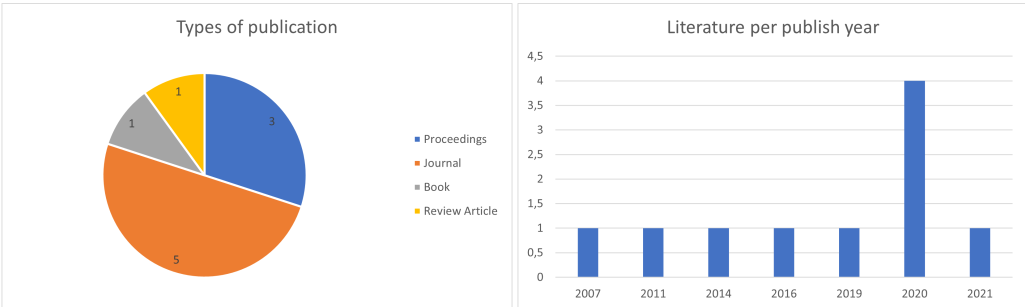
2.3 Results

2.3.1 Analysis

For analyzing the data collected through this study, tools for visualizing and comparing data in Microsoft Excel has been used. Graphs and visualizations created with this program are used to display results regarding the focus areas defined for this literature study.

2.3.2 Focus A: Technologies to enable remote user studies

Concerning focus A, there was a somewhat limited amount of directly relevant literature. However, several articles were indirectly or partially relevant. These articles included some critical findings and knowledge that I could take advantage of when creating the remote user study system.[6]



(a) Literature grouped by publication type

(b) Publish year of literature included

Figure 1: Focus A: Literature overview

As section 2.3.2 displays, most of the literature that was included is from recent years. There are also some older articles, like the study from M. S. Andreasen et al. from 2007 [22]. Although the literature is some years older than the others, it was justified by the article's high relevancy due to directly investigating the topic of how remote user studies can be conducted.

Data Representations

In any type of *remote monitoring* system, either for industrial usage, medical use-cases, or in a user study environment; there need to be decisions made as to how to represent data in a useful manner. This is important as a large goal of these systems is to give relevant information to the *observers* using this system to investigate the data collected from the remote monitoring.

There are many ways that data can be represented in these systems, and it will vary based on the specific use-case of the system. The study from O. Postolache et al. [17] describes a remote monitoring system for a rehabilitation scenario. In their specific system, they collected data about patients performance from the VR gameplay. Data they collected would be things such as score, and how much the user actively used their left- and right-hand to achieve certain things in their gameplay. Additionally, notes could be written and linked to each session as a qualitative *report* about perceived performance

[6] The full spreadsheet including all articles and graphs created for Focus A can be reached on https://docs.google.com/spreadsheets/d/1Q6V3K3mmAlwlRuSdg0v-x7N4bqCgTGzQ14rkJEHUfvs/edit?usp=sharing

or emotions connected to the specific activity the user performed. Their study further states that a system like theirs could facilitate physicians and physical therapists with the necessary information about training outcome, such that they can personalize exercises better for their patients. In turn, this could make patients achieve better rehabilitation results in a shorter time frame. [17]

Another medical application that utilized these graphs to display metrics over time, was a study from A. Choi et al. [18] In their study, they explored how to monitor sleep and respiration remotely. This could deem useful for elderly living alone or people living in places where it is harder to provide medical services in person. In their system, they had different real-time graphs that would display their respiration rate as well as give an indication of what type of sleep state users were in. Their metrical graphs had two different settings; one to display real-time data, and one to display summaries of daily data. Their findings indicated that their system was useful mainly for sleep health management for individuals, but that it potentially also could be utilized to detect acute disease such as respiratory failure. [18]

In the articles investigated in this literature review, some more common paths were taken concerning how the systems would display data. The proposed systems found through the literature review mostly leaned towards either having some sort of graph display of data, either through a metric-over-time type of graph [18, 25], or with additionally having a session-by-session based graph [17]. However, as seen in fig. 2, another common data representation was some sort of graphical or video *replay* of the activity that was being monitored. [19, 20, 23]

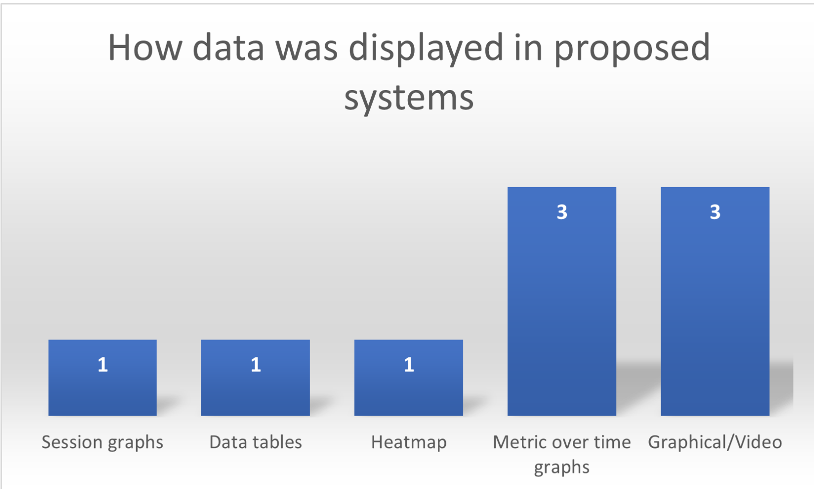


Figure 2: Focus A: Data display used in remote monitoring systems

The study from E. Tsekleves et al. recorded the movement from a patient undergoing rehabilitation and displayed this data as a simulated replay to the therapists [23]. Patients' gameplay was done through a Nintendo Wii remote controller, which also recorded their movements and could be displayed at the therapists' computers. Through using a combination of the Wii remote's acceleration and gyroscope data to track movement, they were able to get 6 degrees of freedom. In turn, this made more accurate representations with their replays. By creating this replay that could be replayed over again, it was seen as highly beneficial for physiotherapists to utilize. With the replay, they could see precise progress, provide good assessment and interact well with their patients, without having to be in the same room as the patients were doing exercises [23]. This brings us well over to the next topic to explore, which is how to run experiments without a study facilitator physically on-site, as well as the issues and benefits of this methodology.

Experiments without Study Facilitator

In the study from M. S. Andreasen et al. [22] they did a systematic experimental comparison of different methods for remote usability testing and compared it with traditional laboratory-based *think-aloud* tests. In their research experiments, the participants were set to try to find usability problems in a system. Their findings indicated that remote synchronous usability testing (i.e. remote experiments but with direct access to a remote facilitator) produced almost entirely the same outcome as conventional non-remote experiments with a physical facilitator in place. They could not find any differences in how long time it took for participants to finish the experiments either when comparing these two methods.

Although that is good news in regards to the research topic at hand, they did also conduct remote *asynchronous* usability testing (i.e. remote experiments without access to a remote facilitator) which did not produce as promising results. The asynchronous method was more time consuming for the participants to complete, and additionally, they identified fewer usability problems [22]. This finding was somewhat of a concern for me. In the *two-week study experiment*, a decision must be made whether the *ExerIsland* platform and experiment task explanations solely can be enough for participants to complete experiment tasks on their own or not. If the case is that it seemed to be an issue, I would have to have some sort of remote presence when experiments are conducted. This would be a much more time-consuming process for me when conducting experiments.

On another note, the present study from M. S. Andreasen et al. did find some other positive effects by conducting remote user studies. One example is that their participants found the study methodology without an on-site facilitator to be less stressful.

> *"I liked this test method better than the traditional method where the test leader looks over your shoulder."*
>
> *- Participant from the study by M. S. Andreasen et al.* [22]

The study by Andreasen et al. [22] presents some of the trade-offs that are present when using a remote user study platform to conduct experiments. As this study seemed to display that remote user studies can be seen as less stressful for participants seems promising for the *remote user study* platform to be developed. On the other hand, the fact that the asynchronous remote user study was more time-consuming and less precise in finding usability flaws indicate that this could potentially be an issue for the *ExerIsland* platform as well. However, one important thing to note is that the context of the two-week study experiment I will conduct is not aimed at discovering usability flaws within the ExerIsland platform. Rather, it will investigate the effects of the gamifications and platform presented. This context does encourage more playful interaction with the platform for participants, compared to a system usability study where the context is to discover and look for issues in a system. Having said that, it will be an important finding to keep in mind when creating the ExerIsland platform, and when evaluating the findings in the practical user study.

Technological Possibilities

Regarding interesting technological possibilities displayed within the current literature, there are a few systems to highlight.

D. Lagun and E. Agichtein [21] developed software to conduct large-scale remote user studies to investigate web-search examination and interactions. Their software, which they called *ViewSer* had a quite simple, yet effective solution. They restricted the user's viewport in the search engine by blurring everything around the mouse pointer position. They displayed that conducting user studies with their remote solution, the *ViewSer* software, induced similar behaviour and results as physical eye-tracking technology did. Due to these results, the system displayed capable of quickly enabling large-scale remote user studies within the topic of web search studies for a much lower cost than in-lab studies.

The study from D. Lagun and E. Agichtein [21] is not directly relevant to the topic of this thesis, which is to create an exergame platform for remote user studies. Yet, it does display that simple, but creative solutions can enable the possibility of running remote user studies for topics that previously have seemed difficult. Moreover, creative solutions can positively affect the ability to create larger-scale studies in shorter time-frames.

Another interesting study is the one by Y. Nishiyama et al. [25], where they created a crowdsensing application for iOS. Their application is an extension of the AWARE framework[7], which is an application created for conducting user studies that can be done through mobile with ease [55]. In their study, they display how parts of their platform worked internally to easily allow participants to join user studies and provide data both from the sensors of their mobile phone and also through answering questionnaires within the application.

In their application, they made it quite simple for participants to join a study; participants could input an *ID* linked to a study going on, and they would be signed up and sending data as a participant. This method allowed users to quickly be able to join user studies within the same platform and seemed to be simple to use for participants. The application provided a dashboard where they could see data they were logging, as well as provide questionnaires for the participants to respond to. In turn, this enabled the possibility to provide some qualitative data or notes from the participants in studies. [25]

The generic user-study possibilities presented by the AWARE framework seemed very interesting to me. Although the system is developed for mobile, the architecture and idea behind the framework seem very similar to the ideas I had for the *ExerIsland* platform. The main difference being that the ExerIsland platform will not be aimed at crowdsensing with mobile, but rather on creating a system that enables remote user studies to be created and conducted with ease.

Ethical Concerns

A review article by M. Madary and T. K. Metzinger [24] sheds light on some of the ethical concerns and best practices that need to be considered when conducting research experiments with VR technology.

[7] awareframework.com

One issue they highlight is that when conducting research with VR in an *online study* scenario, is that there is a higher likelihood that participants may consent to have data collected from them when using the app, but that they become unaware of the continued data collection as they get immersed within the VR context [24]. When conducting user studies, you want the participants to consent and know that their data from application usage will be logged and eventually used in the research analysis. The guidelines by M. Madary and T. K. Metzinger suggest that it should be made very clear to participants, that data from their usage within the application will be logged and used for data collection for the study [24]. For me, this will be important to notice about to participants in the two-week study experiment. Participants should feel informed and are consenting to participate with in-app data being logged.

Furthermore, their review article also highlights that all experiments with VR should follow the principle of non-maleficence:

> *No experiment should be conducted using virtual reality with the foreseeable consequence that it will cause serious or lasting harm to a subject.*
>
> *- M. Madary and T. K. Metzinger* [24]

For the context of the *ExerIsland* platform, I find it important to design and develop a platform that in no way can directly or indirectly be harmful to participants in the study. This will be especially important to concern when developing the VR platform and minigames, as a potential issue is that users may become spatially unaware of their surroundings when immersed in VR. Moreover, this can cause players of these games to cause physical injury to themselves by accident in their homes. Although there is a responsibility on the respective user of the VR application to not become spatially unaware of their own surroundings, it is also crucial that developers and designers of VR applications create their systems in such a way that they do not induce these dangers on their own [56]. For the *ExerIsland* platform, it will be important to not make any of the minigames induce undesirable behaviour from players which may lead to physical injury and will be of vital importance to concern during development. Additionally, I think the informed consent for the VR experiment should contain information that VR may induce certain physiological and psychological effects such as stress or fatigue, as suggested by the review article [24].

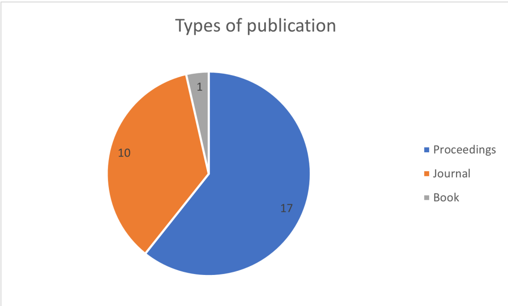
A last important notice from the review article by M. Madary and T. K. Metzinger is that *therapeutic misconception* could be present in VR studies, in a way it could give false hope similarly to how *placebo effects* work. [57]

> *Patients may believe that treatment using VR is better than traditional interventions merely due to the fact that it is a new technology, or an experimental application of existing technology.*
>
> *- M. Madary and T. K. Metzinger* [24]

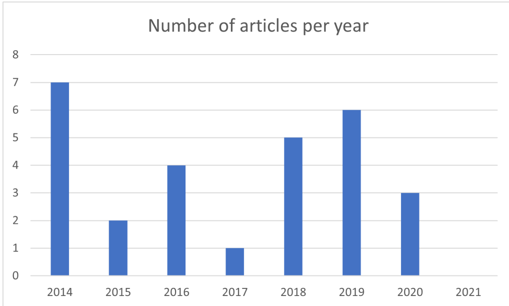
Simply put; it will be important for me to acknowledge this potential bias from participants when they provide their subjective data in the experiments for this thesis. Whilst it likely will be hard to identify whether this effect happens or not, it is an important note to keep in mind when investigating the results of the study experiments conducted.

2.3.3 Focus B: Gamification techniques for exercise motivation

Considering the vast amount of literature out there in regards to rehabilitation and exercise, it can be important to get an understanding of the context of the articles found in this specific literature review. A majority of the publications found were articles from conference proceedings (17/28), as can be seen in fig. 3a below. [8]



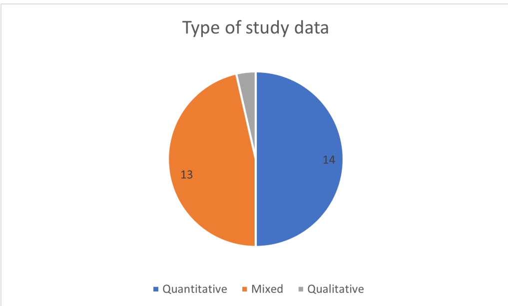
(a) Articles grouped by publication type



(b) Number of articles per year

Figure 3

Additionally, fig. 3b displays that there is some variance in terms of what year the different articles got published. One thing to note is that no articles published in 2021 were included in the final list of literature.



(a) Type of study data in articles

(b) Type of research approach

Figure 4

A large majority of the literature is experiment-based (see fig. 4b). There do exist many different types of studies out there on this research topic. However, the reason this review contains so many experiment-based articles is that I was actively trying to retrieve articles that had done practical research themselves in the search strategy (see section 2.2.2). The reason behind this was that I was expecting practical studies performed on participants to be more relevant for the focus area. On the other hand, this may also cause a bias in the studies found which may give a different picture than another approach would do. Only *one* article was case-study based [41], and *one* was survey-based [27]. The survey-based study was classified as survey-based because their main data collection was through surveys, but even this study was also in fact very experiment-based.

[8] The full spreadsheet including all articles and graphs created for Focus B can be reached on https://docs.google.com/spreadsheets/d/1oGQFIvWqajvStkuM9JoYW2k25QNP3rVaXBpjGxf0kh8/edit?usp=sharing

Figure 4a displays the different types of study data collected in all studies. All studies had some sort of quantitative data, and 13 of the studies also had some qualitative data, putting them in the *mixed* category. However, there were no studies that were exclusively based on qualitative data.

Another note which is important to notice before answering evaluating the results in regards to *focus B*, is that not all of the included articles were focused on specifically *exercise* motivation. Some articles had a main objective of investigating for example *task performance* and motivation to improve in that task. As these studies also investigated gamification elements and software interventions as a means to motivate someone to engage in an activity, these articles were also included. The different research topics in the articles can be seen in fig. 5.

Main Research Topics

There are mainly two research topics that were the most prevalent in the studies found (see fig. 5). One of the topics was *exercise motivation*, measuring whether an intervention made participants perform a specific task better or worse. The second main topic was *task performance*, focusing on whether peoples' motivations towards exercise or physical activity patterns changed after introducing an intervention. Although not all studies main topic was *exercise* motivation specifically, all studies had either a primary or secondary objective of looking at motivation, which is likely to be transferable when applied to an exercise setting as well. I also found that there were several *research and development* studies out there, that conducted *usability* studies on their implemented solutions or platforms. A total of 5 studies had a larger part of their topic being to investigate the usability of their suggested solutions.

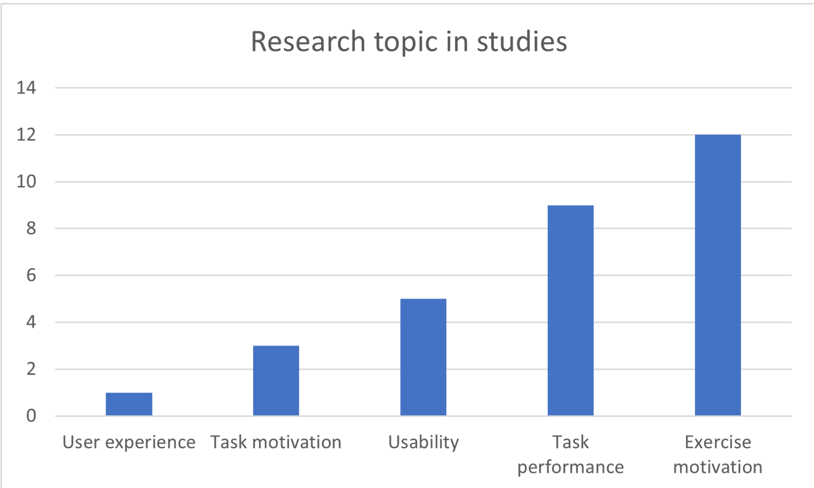


Figure 5: Research topic in studies

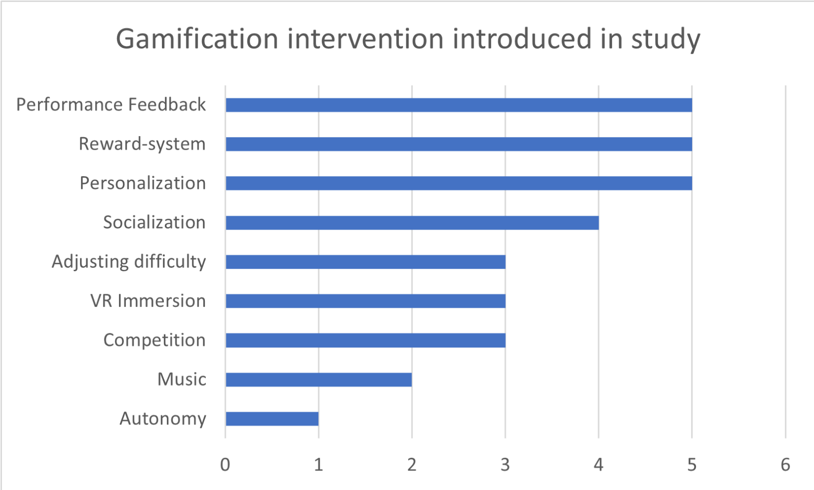


Figure 6: Gamification interventions introduced in studies

Figure 6 displays the different main gamification interventions that were introduced in the studies. Some studies had one main intervention introduced, and some had several. The only thing which is filtered out from this graph is the two categories *Exergames* and *Game*. Some papers did not specifically emphasize what gamification elements they had introduced. Though, they specifically stated that they created an exergame or gamification of an activity with specific goals or actions users could do. This led to these categories being put on almost all studies and would give a wrong perspective as they do not say much about specific gamification interventions except for the fact that they were using games in their intervention.

The three most prominent gamification elements were *Performance Feedback*, *Reward-system* and *Personalization*. Performance feedback in this context was implemented in studies in ways where they were giving real-time specific feedback to the user of whether they were performing actions correct or what they could improve upon [33, 45, 49–51].

Reward-systems on the other hand are more in the ways of rewarding the users. Either through playing certain sounds which are meant as *celebration* when completing things in the games. As well as through giving the user scores or virtual trophies which could be unlocked in the respective games [29, 38, 39, 44, 50].

Personalization, in this context, was to adjust the experience or exercise accordingly to what the perceived preference of the user would be. This could be done through the means of an algorithm or machine-learning. These studies used input from the game activities by the users or data that the users had input themselves. This data would potentially say something about the personality trait of the person, or general preferences in the presented activities. In turn, this would personalize the experience to become more enjoyable and create more engagement for the user. [28, 29, 33, 35, 52]

Some of these results found here, align well with the results in J. Koivisto and J. Hamari's literature review [16]. Their results also indicated that the most occurring gamification intervention was *scoring* as well as *goals*. *Scoring* can be comparable to the *Reward-system* classification defined in this study. While there is no classification specifically defined for *goals* in the classification in this literature review, it can be comparable to the games which got the *Game/Exergame* classification stamp. Almost all of these games did contain some sort of *goal* for the user to complete, and the present study displays a total of 21 studies classified as either of these two.

In general, the gamification interventions indicate some trends of the most common paths that are taken by researchers when it comes to implementing a gamified intervention. It is hard to tell whether this is because these top-trending interventions are the most efficient in motivating users to perform certain activities or if it is due to other reasons. It could also be that these gamification elements are the simplest to implement in practice from a software engineering perspective. Meaning, it is not possible to say for sure if these decisions are purely based on selecting the most interesting gamification element to research, or choosing what is easiest to implement for the software developers.

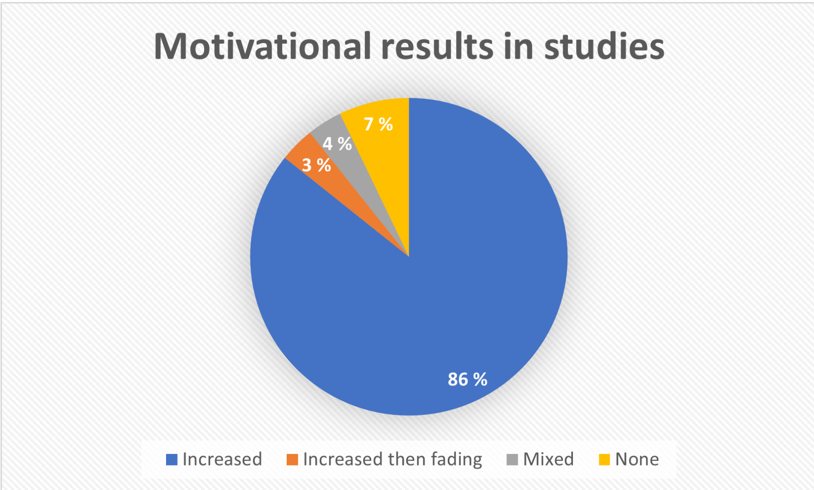


Figure 7: Motivational results in studies

Figure 7 indicates that almost all studies were showing results in a positive direction when it comes to effects on subjects' motivation. In turn, this makes it even harder to interpret whether any specific gamification element is chosen because they prove to have better motivational results for participants, as it seems all the gamification interventions seem to provide something motivational.

In summary, it is hard to get an indication of shared reasoning as to why specific gamification elements are more prominent than others. It is reasonable to see that some gamification elements are easier to implement from a software engineering perspective, e.g. a reward system is easier to implement than a dynamical difficulty-adjustment system. However, in regards to *focus B;* the results indicate that several different game elements are often reoccurring in many studies as seen in fig. 6. The findings of this literature study indicate that the most occurring ones are *performance feedback*, and *reward-system*, tightly followed up by *personalization* and *socialization*.

Technologies Utilized in Research Studies

In regards to investigating the technologies utilized in research studies for focus B, I also labeled technologies that the studies utilized.

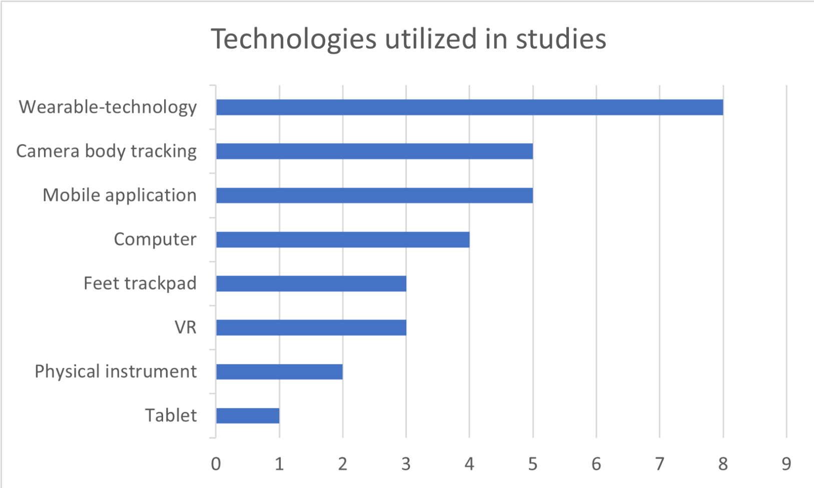


Figure 8: Technologies utilized in studies

Figure 8 displays that certain technologies are more popular than other ones. The most utilized technology is *wearable-technology*, followed by *mobile applications* and *camera body tracking*. *Wearable-technology* is very relevant in professional healthcare settings as they can provide active monitoring of specific body function metrics, e.g. heart rate, blood pressure, or motion tracking. Wearable-technology has also become more popular in general health and wellness, especially with the increased popularity in Internet-of-Things (IoT) in recent years. Very many applications have been developed for smartwatch technology, and they are a cheap and simple way of measuring things such as movement or heart rate for people performing exercise.

For a very long time, *Camera body tracking* systems have been utilized in the creation of exergames. This is because it is a simple, non-intrusive way of interacting with videogames, letting the user just use their body in front of a camera without having to wear any special equipment. Prior to recent improvements in fully immersive VR technologies such as HTC Vive [58] or Oculus [59], one of the main technologies used in many studies were camera body tracking technologies such as Microsoft's Kinect. This still shows in the present results that other technologies are seemingly increasing in popularity, but camera body tracking and wearables seem to be the most dominant technologies used for physical activity research purposes currently.

The least popular technologies found, were *tablets* and *physical instruments*. Tablets, can to a large degree be seen as similar to *mobile application*, while also different to some degree. I think that tablets can be used as a good tool to create serious games on, as they provide easy gesture interactions with big touch screens. These can give some technological advantages when it comes to ease-of-use, especially for the elderly which also might have vision impairments. However, in daily use-case scenarios, tablets are rather inaccessible. They are large and not easy to bring everywhere, compared to their smartphone counterpart.

Physical instrument on the other hand, is in this context that someone has created controllers for an exergame with an existing physical instrument. In the study from G. Osorio [51] this was utilized. They used a cycling exercise machine and made it into a controller for an exergame with visual feedback. This is an interesting approach that lets users work with *controllers* they are familiar with while interacting in a gamified environment.

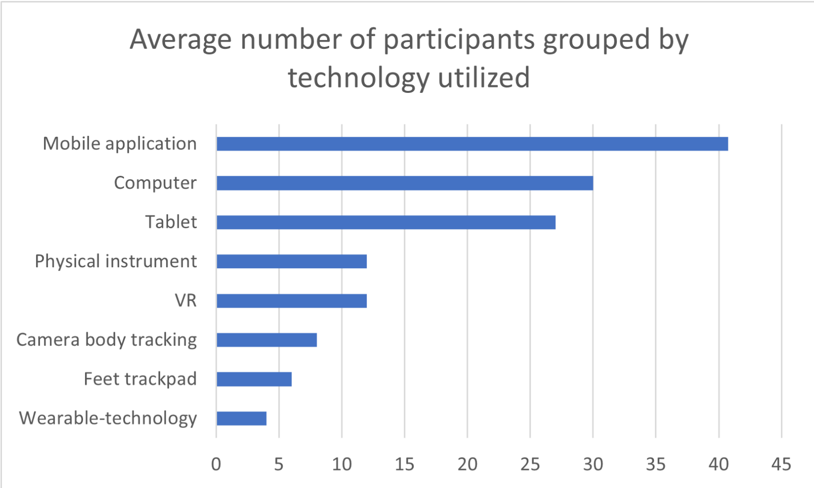


Figure 9: Average number of participants grouped by technology utilized

Looking at fig. 9, there seems to be an indication that certain types of technologies more easily can have a higher participation rate. *Mobile application* seem to have the highest average number of participants, while games just made for *computers* has the second-highest rate. This might be because it could be easier to query for subjects to participate with these technologies. After all, many people already have this hardware in their homes. They are already familiar with them, thus they can easily join these studies. Several of the other technologies have fairly high entry-cost and seem to be harder to conduct with many participants, as can be seen with for example *VR*.

To summarize in respect to technologies utilized in present studies: Some specific technology directions seem more popular than others as seen in fig. 8, but at the same time there is a large range of different technologies being utilized in studies. The two most popular ones are *wearable-technology* and *camera body tracking*. These are closely followed by *mobile applications* and *computers* which seem to give a higher participation rate in studies.

Long-Term Motivational Results

I tried to investigate whether any of the studies displayed any long-term motivational results. From the sample of studies gathered through this literature review, very few studies lasted long enough to be able to measure whether they had any long-term effects or not. While most studies have positive results in regards to motivation (see fig. 7), it is not necessarily true that these effects are transferable to the long-term.

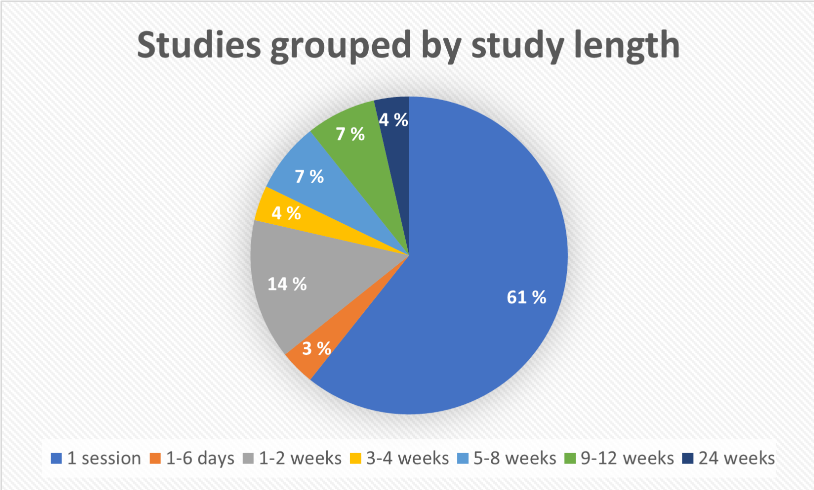


Figure 10: Studies grouped by study length

Figure 10 displays that 61% of the studies included was only conducted through 1 single session, which can not really give any indication of long-term engagement or motivation in the system. The fact that many studies tend to not test their intervention over a longer period is a very huge drawback. From an experiment standpoint, it is easy to understand how it can be simpler to conduct experiments in single sessions. Yet, they can only indicate short-term improvements, and those are not necessarily transferable to the long-term. Moreover, fig. 10 displays that only 22%, or 6 studies lasted 3 weeks or longer, with the longest one lasting 24 weeks. It is hard to classify exactly how long something must be motivating until it is considered as a long-term motivational intervention. However, I did highlight three different categories of *long-term* motivation before conducting the literature review:

3 months or longer: Indicates that there are some parts of this system that either extrinsically motivates the user to persist in using it, or has been able to intrinsically motivate the user. Intrinsic motivation could potentially indicate long-term motivation, but it is not easy to measure whether the motivation is intrinsic or extrinsic by rewards from the gamification.

6 months or longer: Similar as above, but at this point, it is fair to say that the user most likely also have some sort of intrinsic motivation to continue performing the said activity. The user is most likely no longer solely being motivated by extrinsic rewards.

12 months or longer: In my perspective, I think that if a user is motivated to perform said activity for 12 months or longer, that would in itself indicate a long-term motivational result.

None: Is true if none of the above applies, because it is hard to infer if short-term results actually lead to anything in the long-term.

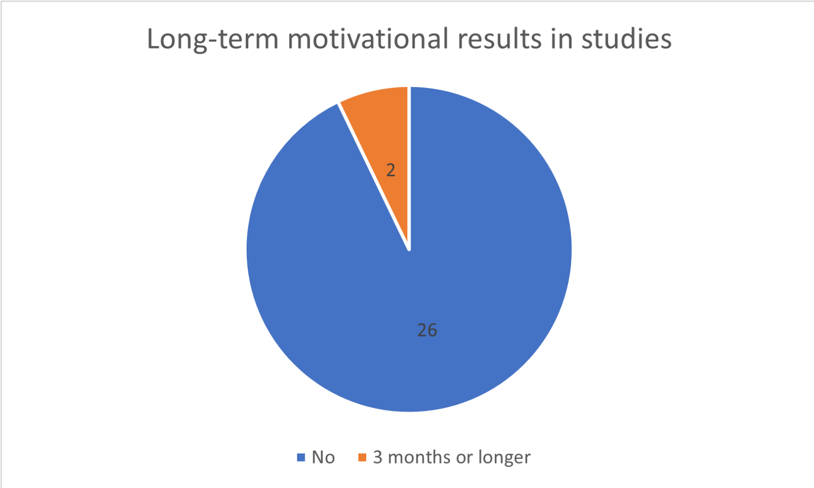


Figure 11: Long-term motivational results in studies

As described earlier, very few of the studies were conducted in the long term, meaning that there would be very few results being able to conclude in any long-term motivational results. Having said that, there were *two* studies which did show motivational results that lasted 3 months or longer (see fig. 11).

The study from A. Wilkinson et al. [35] proved continued motivational engagement over a three-month period with the usage of *Tablet* technology and a personalized serious game for elderly patients.

Secondly, the study from A. R. Gonçalves et al. [47] also showed maintained motivation over three months by using *camera body tracking* as their technology and *music* as one of their key interventions in their exergame.

Another study that is interesting to mention is the study from A. Jaume-i-Capó et al. [28]. This was the longest-lasting study (24 weeks). They also used *camera body tracking* as their technology intervention and focused on *personalization* in their exergame made for people with cerebral palsy. However, throughout their 24-week study, their results indicated that motivation was increasing early on in the study. Nonetheless, later on, that motivation faded away. Participants became reluctant to keep on with performing the said activities. It is important to notice that very many of these different games will be implemented and polished to different degrees, either making them more or less prone to be long-term efficient ways to engage users.

Yet, it is important to observe this specific study, as it was the longest-lasting one. It does indicate results that could be a potential trend with several of the gamifications in the studies explored. It could very likely be that the other studies simply did not last long enough to see that their intervention might not be that motivational in the longer term.

To conclude the *long-term motivational findings*, there is not really enough data provided in the literature at hand to be able to draw any conclusions on what technological interventions seem to keep motivations in the longer term. However, the results indicate that certain interventions have proved to give some motivational results in relatively longer periods, but it is hard to conclude how long this motivation potentially could last.

Other Findings

In addition to the results listed above, there were some other findings that could be relevant for the overall picture of the study topic at hand.

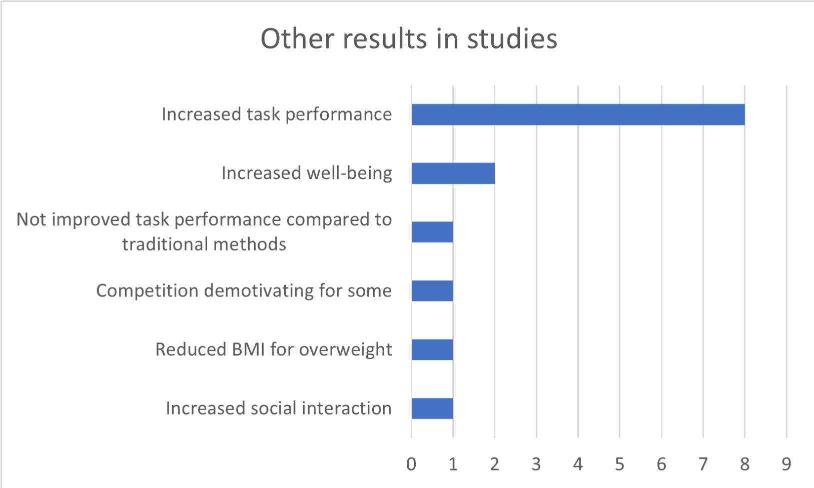


Figure 12: Other results in studies

Figure 12 displays some of these other findings from the literature. The *increased task performance* is the most occurring one. This was often the main or secondary research objective in many of the studies included. However, there are also some other interesting findings here, for example, the study of D. H. Goh and K. Razikin [44] indicate that certain participants were feeling demotivated by competitional aspects introduced in their exergames. These participants often stated that the fact that they could see that they were underperforming compared to other people, made them feel worse about themselves and less motivated to continue [44].

Another finding is that A. R. Gonçalves et al. [47] found that their gamified intervention was not making participants perform better than with traditional exercise methodology, despite their findings in regards to increased motivation in a three-month period.

In the study from Z. Zhao et al. [52] they tried to model their users based on some data about their personality traits. With this data, they tried to use a machine-learning model to suggest *physical activities* that the users would be interested in doing. Although showing a somewhat useful system, their findings indicated that personality traits alone are not enough to suggest very valuable types of physical activities for their users. Meaning, that their personality only has a limited influence on their preferences in physical exercise, and more data should be considered for personalization use-cases. [52]

Overall, secondary results that were found through the studies are mostly indicating positive effects in addition to the motivation measurement. *Increased task performance* being the most noteworthy one.

2.4 Limitations

In all work conducted, there should be room for critique. In this specific context, it can shed light on potential flaws of the systematic literature review conducted. There are a few points I would have wanted to improve if I were to do the literature review again in the future.

2.4.1 Conducting literature review

I found it hard to conduct a literature review solely on my own. There are two main reasons for this: First of all, the sheer amount of work that is required to properly conduct a good systematic literature review is a bit overwhelming for one single person. Secondly, I think that being only one person in a systematic review like this gives more bias than a team with more people conducting the review. Throughout the project, I had to instinctively trust my personal opinions on different matters. There are several occasions where I believe it would have been advantageous to have the possibility to discuss certain issues or get second opinions on different subjects. This would be very valuable not only for discussing results to draw conclusions to focus areas but also in collecting and classifying the specific papers to utilize.

2.4.2 Focus A: Technologies to enable remote user studies

Limited amount of directly relevant literature

The main limitation to the literature study conducted for focus A was the limited amount of directly relevant literature to this topic area. When I conducted my literature search, most findings were in regards to running in-lab remote experiments with physical or technical equipment, and not with *people* as subjects.

This was a rather large concern, as I spent a considerable amount of time trying to find directly relevant literature on the topic of how to conduct or create a system to enable *remote user studies*. Due to this limitation, two issues arose:

Firstly, the literature that would be included in the literature study would need to be more generic and not that specific in investigating technology around *remote user studies*. Instead, I had to include literature that was somewhat less relevant. The literature still touched upon the topic with some relevant information that could be important to answer the topic being explored. This means that some articles which were found less relevant, are still included.

Secondly, because it was hard to find specific literature that gave concrete answers to the topic area at hand, I did end up with a fairly limited amount of literature for the literature study. With the final literature study containing only *10* different articles. I think that, from the beginning, I probably should have allowed a broader amount of literature to be included in the literature review. Although it could have been literature that was even less relevant, it would have provided more articles to give a broader perspective into certain areas.

2.4.3 Focus B: Gamification techniques for exercise motivation

Too Wide Search Strategy

First of all, the search strings used in this project might have been covering too wide of a topic area, the search strings ended up being quite large (see section 2.2.2) that would include a wide range of topics. They were created to be specified towards motivation, exercise, gamification, or other software engineering technological interventions. Albeit, they were probably a bit too inclusive, as displayed by the number of search results with 1043 total papers also (see section 2.2.2). The execution phase of conducting the systematic literature review could probably be done more systematically if these searches were more narrowing also.

Additionally, when looking at the resulting 28 papers that were used in this systematic literature review, all of them is not specifically targeted to research physical exercise. Although, as explained earlier, motivational results are likely to be transferable when applied to different contexts. However, to give more properly narrowed results for focus area B, it could have been beneficial to have even more literature that specifically was looking at physical exercise as the topic area.

Furthermore, I was initially expecting to see a lot of motivational results in the *'Increased then fading'* category (see Motivation results in table 4). As I thought this was a common issue to many of the less developed exergames out there. Motivation would be increasing early on, but slowly fade away over time as the extrinsic rewards might become less appealing. Concerning the search strategy, I believe that there could have been better search strings, as well as better inclusion and exclusion criteria to include more papers that display long-term experimental research.

Lack of Software Engineering Classifications

One of the key aspects of this thesis is to investigate how software engineering techniques can be applied to increase users' motivation with physical activity. With that, I think there could have been more classifications of software engineering specific categories in the execution phase of this review. For example relating to software architecture choices, development methodologies, or what actors were involved in the development of the different games and software.

These different aspects would shed light on potential new correlations in the data collected through the literature. The main software engineering categories which were classified in the papers were *Technology* and *Intervention*. The intervention was often software engineering centric or a specific gamification element that was introduced in the study. These gave relevant information about the current research topic, and I think that more of these software engineering classifications could give even more interesting findings.

2.5 Summary & Conclusion

As described in section 1.1, technological interventions could be helping with motivating people to exercise and rehabilitate. In fact, it can play a large role in keeping a healthy society in the future. I think this review sheds some light on some of the current states of this research area, and are well complemented by similar reviews [16].

The goals of this literature review were to investigate two focus areas: *Focus A* was about how a system can be made to enable remote user studies, while *focus B* was about investigating the techniques and gamification elements currently being researched in the topic of exercise motivation.

Initially, there was a planning phase that included defining a search strategy and coming up with suiting inclusion and exclusion criteria to find papers relevant to the focus areas. Information from the literature was collected through the execution phase and classified into different categories. In addition, notes about findings were written in the spreadsheets. The total amount of literature articles included in this review was *38*.

Further, the data collected through all of these papers were put together and analyzed in an attempt to give some relevant knowledge to the *focus areas*. For *focus A*, there were several relevant findings from the literature in regards to how remote user studies can be done.

First, the literature displayed some of the different ways data can be represented in these *remote monitoring* systems. Graphical replays seem to be very prevalent in the medical domain, while general data graphs were the most common ones in more casual systems.

Secondly, the literature shed light on some of the trade-offs in regards to conducting user experiments without a study facilitator in place. The literature indicated that participants were feeling less stressed when doing experiment tasks without a physical facilitator in place. On the other hand, an indication was also made that if the participant had no direct access to a facilitator to ask questions to, it could reduce the quality of the data from experiments. [22]

Thirdly, some of the interesting technological possibilities in the space of remote studies were explored. Most interestingly were the ViewSer system that could mimic eye-tracking software to enable large-scale studies quite quickly, [21] and the findings in regards to how the crowd-sensing mobile application AWARE was implemented [25,55].

Lastly, there were several findings in regards to ethical concerns that have to be recognized when conducting user studies with VR and also, remotely. The literature states it is important to make it clear to participants that data from their VR application will be collected and used for research purposes. Further, all experiments with VR must follow the principle of non-maleficence; do no harm. [24]

In regards to *focus B*, it was found that some gamification elements were more popular in studies than others, from the literature collected I found that performance feedback, reward-systems and personalization seemed to be the most popular ones.

Further, it was found that some of the main technologies being utilized in these studies were *wearable-technology*, *camera body tracking* and *mobile applications*. It was also displayed that *mobile applications* and *computer* technology seemed to have a higher participation rate.

In regards to motivational results from the found studies, it was seen that most of the studies were positively oriented in regards to this, often having the participants displaying high engagement with their gamification or technology. However, there were only *2* studies that displayed some sort of long-term engagement in their study. [35, 47]

Overall, this literature review gave me useful knowledge and insights into the topics being researched in this thesis. The information collected from this literature review helped shape the directions of the developed artifacts and experiments conducted in this thesis.

3 ExerIsland: An Exergame Platform for Remote User Studies

In this chapter, I describe the platform utilized to conduct the practical 2-week study experiment. The methodology and technical decisions made for this system, is in many ways directly relevant for *research question 1: How to develop an exergame platform that enables the ability to run remote user studies?*

Note: parts of the design and artifacts described in this chapter is based on my existing work conducted in IMT4807 Integration Project, IMT4307 Introduction to Research into Serious Games and Gamification, and IMT4889 Specialization in Mobile/Wearables. The systems were extended and modified for the context of this thesis.

3.1 System Architecture

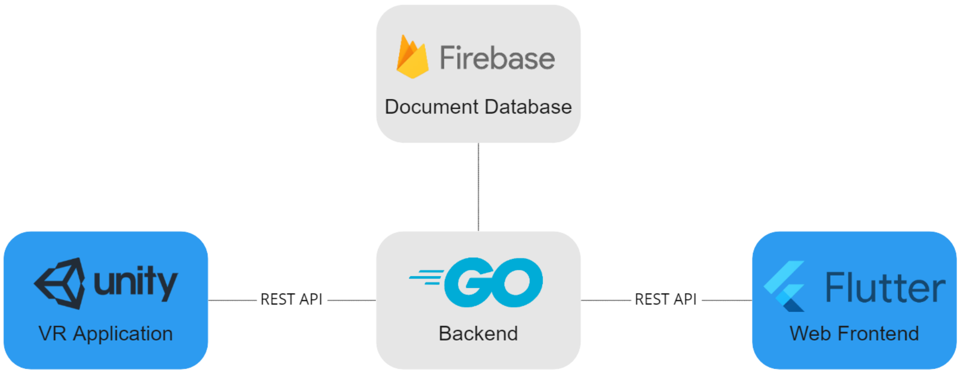


Figure 13: High-Level System Architecture

Figure 13 displays the high-level system architecture of the ExerIsland platform. It consists of *4* main components, with a total of 13097 Source Lines of Code (SLOC) across all 117 code files:

Unity VR Application: The hub containing minigames to play.
Consists of 4091 SLOC across 42 C# scripts.

Golang Backend: Handles data interactions in the platform.
Consists of 2932 SLOC across 18 Golang files.

Flutter Web Frontend: Online dashboard to view data from game activities.
Consists of 6074 SLOC across 57 Dart files.

Firebase Document Database: Storing data for the platform.

The VR application and web frontend communicate with the backend through a RESTful API. For the backend to communicate with the database, it authorizes through the

FirebaseAuth layer with a token that provides administration access to the database. This means that there are no other entities that will directly access the database. All interactions happening on the database will always happen through the backend. The different interactions done with the backend's API is further described in section 3.3.

My development and research process has gone side-by-side. I have used an issue tracker actively for both activities and completed 255 issues throughout the project. The codebase for the developed artifacts are split into two repositories:

Golang Backend & Flutter Frontend:

gitlab.com/akerholten/vr-health-and-wellness-remote-monitoring

VR Application (without proprietary assets):

gitlab.com/akerholten/exerisland-vr-exergame-hub

3.1.1 Choice of Technologies

With respect to choice of technologies, this was done in concern to three main points:

1: What are the capabilities of the technology?

I had to know that the frameworks chosen would be capable of achieving the purpose with the ExerIsland platform. Such that the final platform never was restricted in functionality due to restrictions made by technology.

2: How effective is the technology in regards to rapid prototyping and development?

As I was going to create and implement a quite *intricate* system with a broad set of technologies, it was important for me to know that each component of the ExerIsland platform could quickly be up and running in a Minimum-Viable-Product (MVP)-state.

3: How much time will it take to get comfortable working with the technology?

Similar to the *second* premise, I could not afford to spend too much time on learning a new framework or programming language as it would take away too much time from developing the actual artifacts of the system. This meant that if I were to chose a technology I was unfamiliar with, the time it would take to get confident with this framework would have to be justified by the benefits from the first two premises.

The justification for choosing the technologies for the main components of the ExerIsland systems is described below.

Unity VR Application

For the VR exergame application, there were mainly two options; Unreal Engine[1] and Unity[2]. It was mainly these two options I considered, as they are the most polished and well-developed game engines which are free to use for independent developers. They are also popular in the industry, both for entertainment and more serious purposes.

For me specifically, I knew that both Unreal and Unity possessed the capabilities required to create a well-polished exergame platform for VR. Both provide seamless integration with *SteamVR*[3], which is a tool that promotes rapid prototyping for VR applications.

[1] unrealengine.com
[2] unity.com
[3] steamvr.com

However, I was most familiar with the Unity game engine. As both technologies provide similar tool-sets to create the exergame platform, I decided to stick with Unity for the VR application.

Golang Backend

For backend programming language, there were mainly three options considered: Node.js, C#, and Golang.

For this choice, I saw the different languages as similar in regards to their technical capabilities. However, the final decision landed on Golang as I saw my experience with the platform as very beneficial to allow rapid development in the project.

Firebase Document Database

With concern to the database choice, there were mainly three options I considered; Firebase Realtime Database[4], Cloud Firestore[5] and MongoDB[6].

The technological capabilities of all these systems are very similar, and I had some experience within all three different options. However, as the backend landed on being based on *Golang*, it seemed like Firebase Realtime Database was the most well-documented option to use in combination with a Golang backend. Google also had a well-created Software Development Kit (SDK) [60] package for rapidly connecting and integrating Firebase functionality into Golang backends. Due to the mentioned reasons, I decided to go with the *Firebase Realtime Database*.

Flutter Web Frontend

For the frontend, there were mainly two options being considered; *React*[7] and *Flutter*[8].

In respect to this decision, I deemed both technologies similar in their technical capabilities. However, React is more well-known and documented as it has been an industry standard for cross-platform development for many years. Further, I did not have any prior experience with Flutter or Dart, and were more experienced with React.

On another note, I feel there are some quirks and pain points related to using React and Typescript for web development. Additionally, Flutter seemed to be incredibly well implemented to allow rapid development and prototyping. Flutter has a feature-rich development API that provides native performance. In addition, it contains a lot of UI components and the well-known material design out of the box. [61]

For this specific technology choice, I believed that the speed of development introduced by Flutter would justify the time it would take to learn this new technology. After all, the Flutter framework and Dart programming language seemed quite simple to use. It provides similar features as other web-development frameworks, such as re-usable UI components and hot-reloading during run-time debugging.

[4] firebase.google.com/docs/database
[5] firebase.google.com/docs/firestore
[6] mongodb.com
[7] reactjs.org
[8] flutter.dev

3.2 Data Structure

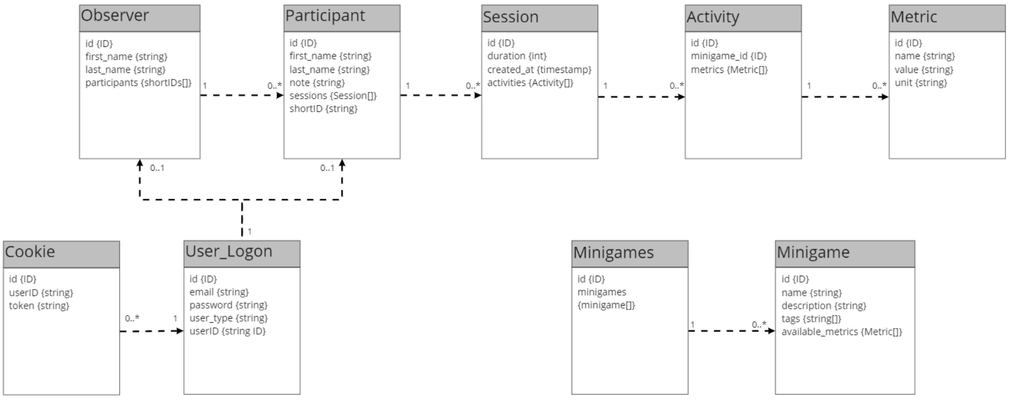


Figure 14: Data Structure

Figure 14 displays the overall data structure from the data stored and used in the application. The data structure is not very complex, but to give context to the different entities stored in the system, these are described below.

Although the database is document-based, I found it simpler to display the data structure through a UML diagram as seen here. It displays the system similar to how a relational database would function in practice. Nevertheless, I find this way of displaying data structures easier to understand, which is why I have done it here.

3.2.1 Observer

The *observer* is the user-account for a *researcher* (or *facilitator*) in the system. It works like a moderator account that has a larger overview. The observer has a list of shortIDs linked to participants. The observer is authorized to access data from these participants and can do so in the frontend application.

3.2.2 Participant

The *participant* entity is quite self-explanatory. It is the user account for a participant taking part in a study. The participant also holds a record of logs from game activities in the VR application. This is done through the means of sessions, activities and metrics.

3.2.3 Sessions, Activities & Metrics

Session

A session is mostly a container for *activites* that has taken place in a specific game session. A game session is created when the user boots up the VR application and plays any minigame in the application. The session persists while the user stays in the VR application, and it will hold a list of several *activities* done by the user in the game session. Additionally, it holds the *duration* of activities in the session and a *time-stamp* of when the session took place. This information is not only useful for the *researcher* but also for the participant themselves when looking at their personal dashboard in the frontend. (this dashboard is described in section 3.5.3)

Activity

An activity is also quite self-explanatory. It is a specific activity that has taken place in the VR application by the participant. The activity is linked to a specific minigame-ID (the game played), and it contains several metrics logged during this game activity.

Metric

A metric is a type of data logged during gameplay in the VR application. These metrics could say something about how the participant performed, how long they were doing said activity, or it can be other interesting data, e.g. how much their *arms* were moving during gameplay.

Metrics are logged in the VR application through the use of `MetricLoggers` defined in Unity that logs different types of data. (see section 3.4.6) Any metric contains an ID, name, value and a unit (definition of what type of metric is logged, e.g. seconds or meters). An example of a Session-object, with its activities and metrics, can be seen in listing 3.1.

```json
{
  "id": 1,
  "duration": "211",
  "createdAt": "2021-04-15T19:00:08Z",
  "activities": [
    {
      "minigameID": "ReactionTimeTrainer_Minigame",
      "metrics": [
        {
          "id": "Duration",
          "name": "Duration",
          "value": 102,
          "unit": "seconds"
        },
        {
          "id": "Average_ReactionTime",
          "name": "Average reaction time",
          "value": 759,
          "unit": "ms"
        },
        {
          "id": "Score",
          "name": "Score",
          "value": 1132,
          "unit": "score"
        },
        {
          "id": "Arm_Movement",
          "name": "Arm movement",
          "value": 82,
          "unit": "meters"
        },
        {
          "id": "Calories_Burned",
          "name": "Calories burned",
          "value": 7,
          "unit": "calories"
        }
      ]
    },
    {
      "minigameID": "DroneShooter_Minigame",
      "metrics": [
        {
          "id": "Score",
          "name": "Score",
          "value": 7575,
          "unit": "score"
        },
        {
          "id": "Arm_Movement",
          "name": "Arm movement",
          "value": 91,
          "unit": "meters"
        },
        {
          "id": "Calories_Burned",
          "name": "Calories burned",
          "value": 7,
          "unit": "calories"
        },
        {
          "id": "Duration",
          "name": "Duration",
          "value": 109,
          "unit": "seconds"
        }
      ]
    }
  ]
}
```


Listing 3.1: JSON object of a Session containing Activities and Metrics

3.2.4 Minigames

A list of minigames is also stored in the data structure. However, it is not stored in the database, but in a hardcoded list on the Golang backend. The purpose of this list is simply to have all the metadata about minigames be stored in one singular place, as it is retrieved to display some information on the frontend. An example of minigame metadata stored on the backend is displayed in listing 3.2 below.

```
Minigame{
        Id:          "ReactionTimeTrainer_Minigame",
        Name:        "Reaction Time Trainer",
        Description: "Based on BATAK reaction time trainer for
            football-keepers, the reaction time trainer minigame focuses
            on providing a series of water bubbles that the player should
            hit when they light up and make a noise.",
        Tags: []string{
                "Physical", "Cognitive", "Arms", "Hand-eye coordination",
                    "Hearing", "Sound",
        },
        AvailableMetrics: []Metric{
                Metric{
                        Id:    "Duration",
                        Name:  "Duration",
                        Value: 0,
                        Unit:  "seconds",
                },
                Metric{
                        Id:    "Arm_Movement",
                        Name:  "Arm movement",
                        Value: 0,
                        Unit:  "meters",
                },
                Metric{
                        Id:    "WaterBubbles_Hit",
                        Name:  "Water bubbles hit",
                        Value: 0,
                        Unit:  "hits",
                },
                Metric{
                        Id:    "Score",
                        Name:  "Score",
                        Value: 0,
                        Unit:  "score",
                },
                Metric{
                        Id:    "Calories_Burned",
                        Name:  "Calories burned",
                        Value: 0,
                        Unit:  "calories",
                },
                Metric{
                        Id:    "Average_ReactionTime",
                        Name:  "Average reaction time",
                        Value: 0,
                        Unit:  "ms",
                },
        },
}
```

Listing 3.2: Example Minigame metadata from Backend

3.2.5 Cookies & User_Logons

The *cookie* and *User_Logon* objects stored in the database are used for authentication and authorization with the application. Any user is authenticated with their User_Logon entry in the database, it contains their credentials with email/username and password which are salted and hashed on the backend. This entry in the database points to either an *observer* or *participant* user with the user-ID.

When users log in, a cookie is stored in their browser. The cookie is used to authorize their activities throughout the session in the frontend application. For any of the interactions the user makes in the application, a cookie must always be present and verifies whether the user is permitted to do specific operations or not.

3.3 Backend API

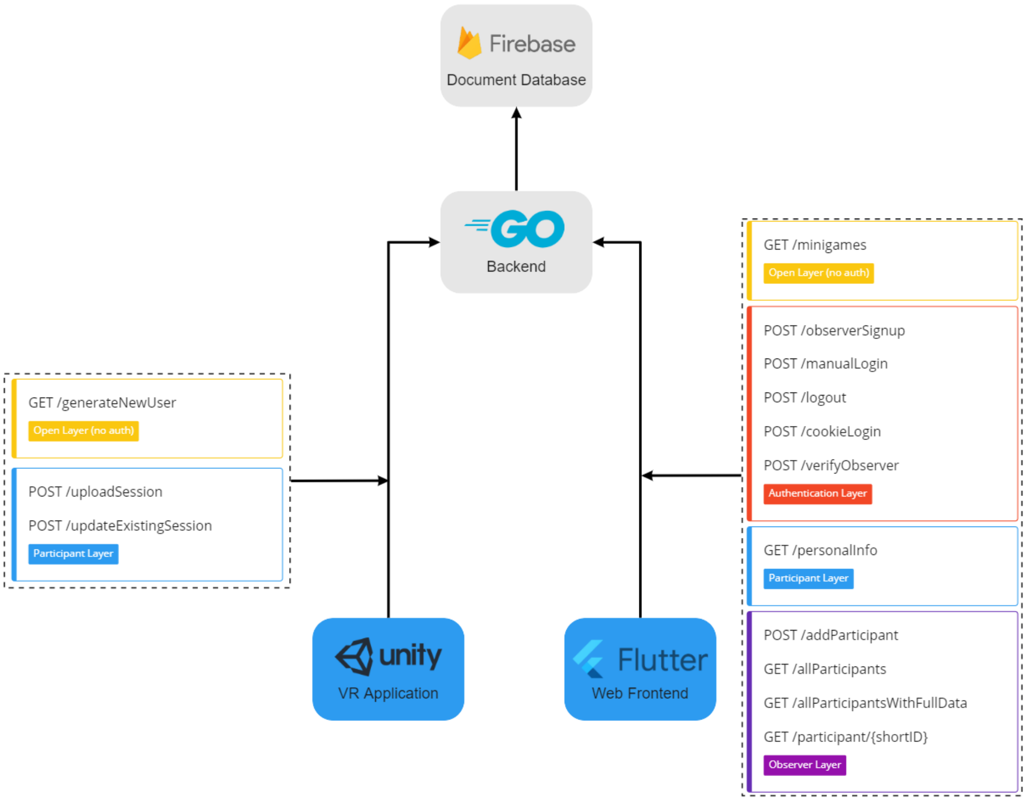


Figure 15: Data Interactions on the backend API

Figure 15 displays the different interactions that can take place in the API of the ExerIsland system. It mainly consists of basic Create, Read, Update and Delete (CRUD) operations interacting with the data models described above in fig. 14. The data flow is divided into two parts; one for the interaction between the VR application to the backend, and one for the API between the frontend and backend application.

3.3.1 VR Application to Backend

From the VR application, there are only two layers in the API, the *Open Layer* and *Participant Layer*. This layer from the VR application does pose some security flaws in the system. However, during development, the decision was made to focus on ease-of-use and rapid prototyping compared to considering security aspects. Especially since the goal of this thesis was to create a prototype system to conduct remote user studies, it was more important to focus on this specific aspect, rather than spend time implementing security parts of the software.

Nonetheless, I acknowledge that it is a weakness of the system. Some of these security flaws are further discussed in section 6.1, where I also try to explain some potential solutions to solve parts of the present security issues.

Open Layer

GET /generateNewUser

This interaction provides the functionality necessary for a user to connect to a study experiment as a participant. In response, the user receives a *participant-ID* which will be stored in the Unity application for logging their activities and be displayed to the participant.

Doing this API call does not require any authorization, which is the way it must behave such that new unregistered participants can sign up for a study.

Participant Layer

POST /uploadSession & POST /updateExistingSession

These two POST operations enable participants ability to log data from their sessions to the database. The user authorizes in a very simple way by using their participant ID when sending the requests.

3.3.2 Web Frontend to Backend

Open Layer

GET /minigames

This interaction is the second and last interaction in the API that does not require any authorization. This simple query to the backend directly replies with a hardcoded list of existing minigames in the system. It is simply a way to store some meta-data about minigames in a single part of the application. This list of minigames is also not stored in the database, but exist as a hardcoded list of minigames on the backend codebase. This was not only done for convenience but also for performance reasons.

In terms of convenience, there was no implemented interface to interact with a list of minigames from an administrative perspective, and it was just simpler to maintain a hardcoded list. Concerning performance, it did reduce unnecessary look-ups in the database from the backend, which was even more important for this specific query, as it was in the *open layer*. Malicious users could potentially misuse this interaction to perform a very impactful Distributed Denial-of-Service (DDoS)-attack, which further emphasized the decision to make this open call less requiring in terms of processing on the backend.

Authentication Layer

POST /observerSignup

POST /observerSignup is called to the backend when a user creates a new account through the frontend application. This user becomes an *observer* user and would be possible to use to monitor participants in a study.

POST /manualLogin

This login interaction is called from the frontend when logging in with username and password. It is used both for *participants* and *observers*.

POST /logout

A simple logout call that removes the cookie from the client's browser, and also deletes the cookie entry in the database.

POST /cookieLogin

This call is used on the frontend to verify that the client's session persists. All the interactions done by the user is authorized through using the cookie. To not have any issues while using the app, the app continuously checks that the client's session is still active through using this cookie verification call.

If the cookie no longer is valid and the session is no longer available, the client will get redirected to the login page of the frontend.

POST /verifyObserver

This is used as an *authorization* to verify whether an account is permitted to access certain parts of the application. Because the application is designed around having two separate dashboards, one for *participants* and one for *observers*.

Participant Layer

GET /personalInfo

Apart from several of the interactions above from the *auth*-layer, this is the only operation the participant will make use of on the frontend. This is a simple call to the backend retrieving all existing personal data logged from sessions in the VR application. This data is displayed through graphs and statistics for the users to view (further described in section 3.5.3).

Observer Layer

The *Observer Layer* is the layer for users authorized as *researcher* (or *facilitator*) in the system. It is called the Observer Layer as it accurately describes what the use-case for these users in the platform is; to *observe* activities of participants in the user studies.

POST /addParticipant

If the observer would like to manually add participants to an experiment or his dashboard, this can be done through this call to the backend.

GET /allParticipants

This call retrieves all participants in the system which is linked to the observer. To mitigate unnecessary performance hits in the backend by retrieving *all* data about all participants through this call, it only retrieves the basic information which is displayed about participants in the observer's dashboard. (the observer's dashboard can be seen in section 3.5.2)

GET /allParticipantsWithFullData

Differently from the simpler call above, this call retrieves *all* data about all participants, containing basic data and also all the logged sessions from the VR application. Although there currently are no parts of the frontend that utilizes this interaction with the backend, it exists there so it can be used if there would be a use-case for it.

The main reason I created this functionality, was to have a way to retrieve a full *JSON*-response containing all the data from all participants throughout the study. This was useful, as this JSON data was utilized for data analysis when the two-week user study was finished.

GET /participant/shortID

This call retrieves the full data about one single participant. In practice, it does the same as the *GET /personalInfo* call from the participant's perspective, but in this case, in the perspective of the *observer*. The observer sends a short-ID, which is similar to the participant-ID to retrieve data about a specific participant. This displays the participant's dashboard to the observer.

3.4 ExerIsland: VR Application

3.4.1 The Main Hub

Figure 16: Island environment of the main hub in VR

The main hub environment of the VR application[9] is designed to be an open and calming atmosphere for users. An island, with an ambient atmosphere including sounds of birds and water washing in on the shoreline. There are many other alternatives to how main hubs can be created in VR, and many existing platforms out there tend to go for a minimalistic sci-fi approach. However, the design idea here was to have a more *welcoming* and *warm* atmosphere, which could be positive to users in an exergame application like this.

With this approach, it could allow for a more interactive hub environment where users could explore and play with different objects in the scene. Although it is important to mention that this was not part of the focus area in this development project, and the hub environment ended up less interactive than initially intended. The main implementation focus in the VR app was on polishing and making minigames feel good to play for users. In addition to implementing the solution that enabled the remote user studies.

Minigame Portals

Minigames are entered through a portal in the minigame portal area. Each minigame has a prefab with objects from the respective game scenes. This serves two purposes: Firstly, it can make players curious about what exists in minigames and make them want to test them out. Secondly, it gives the players some insights into what to expect from the minigame.

Figure 17: Sign pointing towards minigame area

[9]Video footage from the VR application and minigames can be seen at youtu.be/RaJcLMOqtss

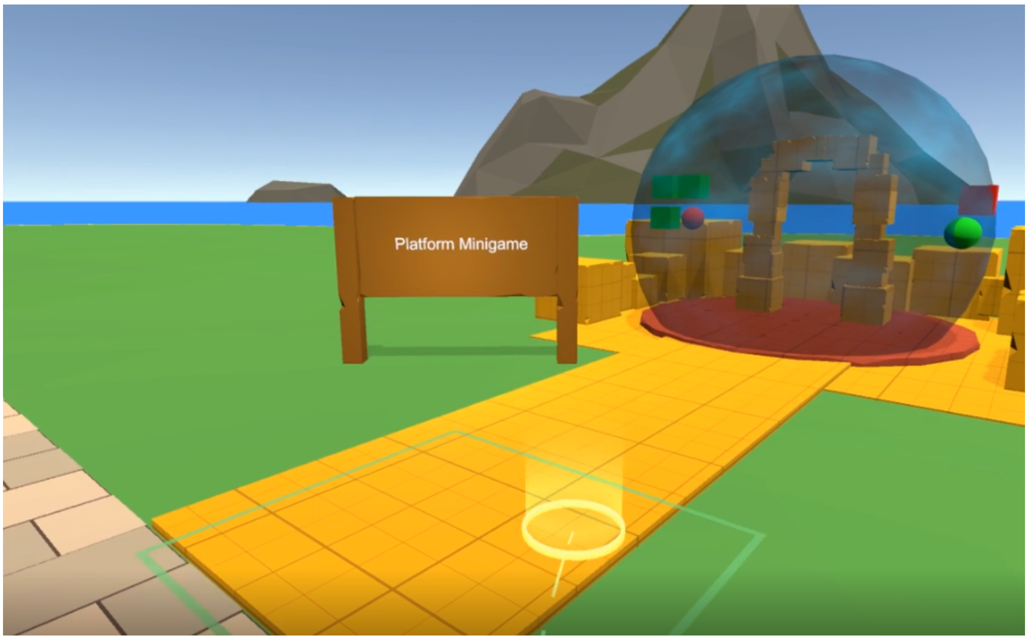


Figure 18: Portal to the Platform Minigame

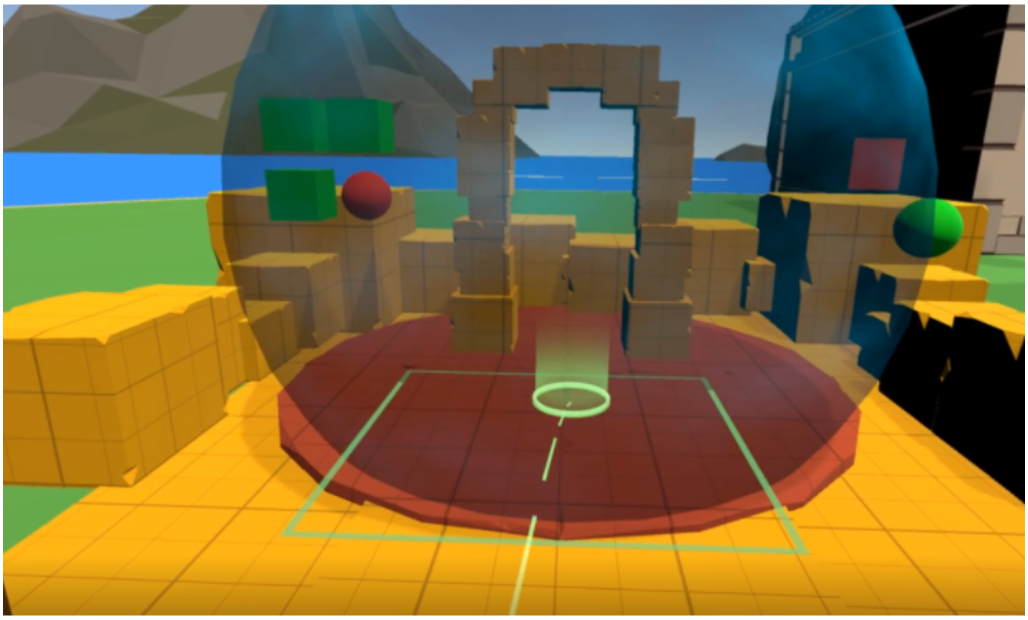


Figure 19: User loads the minigame by entering the portal area

3.4.2 Class Structure Diagram

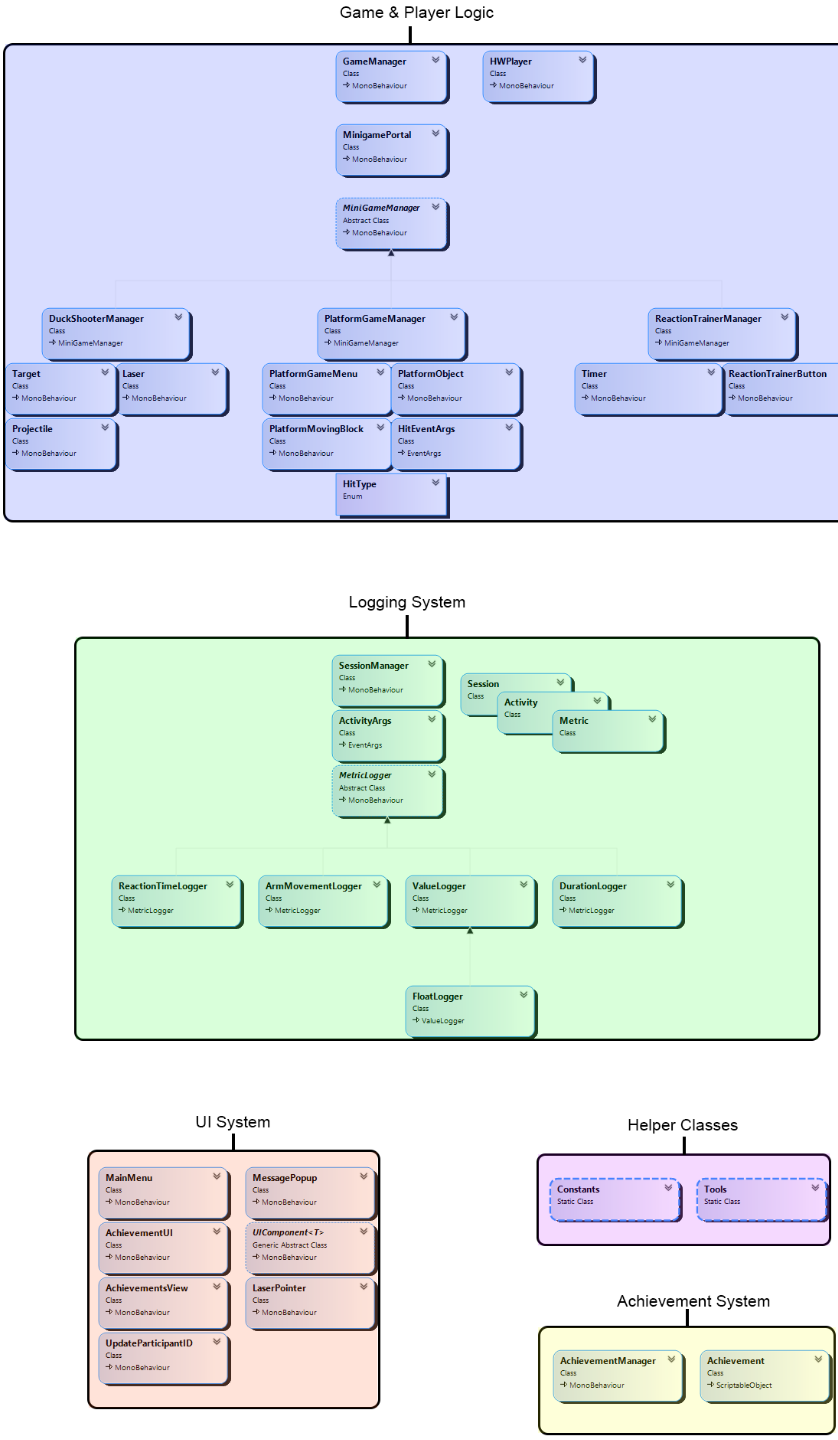


Figure 20: Class Structure Diagram for VR Application

The main functionality of the VR application in Unity is divided into *5* parts:

Game & Player Logic: Contains scripts and classes that handles the player and game logic in the overall VR hub, in addition to separate scripts for the different minigames.

Logging System: Is the system handling all the logging in the minigames. It is responsible for logging the data, storing it in the `SessionManager`, and sending it to the backend with HTTP requests for storing it online.

UI System: The system responsible for all UI functionality in the application, gives the user graphical user interfaces to interact with in VR.

Helper Classes: Shared constant variables and helper functionality shared between several scripts.

Achievement System: The system for in-game achievement tracking.

3.4.3 Minigames

The VR application contains *3* different minigames to play. Each minigame tries to capture different types of atmospheres and engage users in various gameplay. Different gameplay caters to divergent player types, and with minigames designed quite distinctively, it could potentially increase the replay value in the application. The purpose of the minigames is to induce physical activity for the player in some sort as they are designed as exergames.

As seen in fig. 20, each minigame had a `MiniGameManager` script which would be the entity controlling game state in the minigame. As several of these minigames would have similar functionality, they all derive from a `public abstract class MiniGameManager : MonoBehaviour`. Consecutively, this also made the process of creating and adding new minigames to the platform more streamlined, as some of the baseline functionality would be there in the `MiniGameManager`. To some degree, creating a minigame manager would be like following a recipe to quickly have some basic minigame functionality up and running.

Listing 3.3 displays parts of the functionality which is contained in the `MiniGameManager`. Much of the logic contained in the class is relevant to activity logging, gameplay logic, game menu interaction and metadata about the minigame.[10]

[10]The full MiniGameManager.cs script can be seen in gitlab.com/akerholten/exerisland-vr-exergame-hub/-/blob/master/Assets/_Assets/Scripts/Managers/MiniGameManager.cs.

```
// Abstract class that declares interface the minigames must follow each
    minigame manager will deal with independent logic themselves
public abstract class MiniGameManager : MonoBehaviour
{
    [Header("Minigame Manager Info")]
    public int score;
    public int difficultyLevel;

    // liveDifficultyLevel is used for games with scaling difficulty
    protected float liveDifficultyLevel;

    public bool IsPlaying => _playing;
    protected bool _playing = false;

    // gameHasBeenPlayed is true if a round of the minigame has previously
        been played in this session
    protected bool gameHasBeenPlayed = false;

    [SerializeField] protected PlatformGameMenu gameMenu;
    [SerializeField] protected GameObject gamePlayObjects;

    public float GameDuration => Time.realtimeSinceStartup - gameStartTime;

    // gameStartTime is stored when StartGame() is called, used to
        calculated game duration
    protected float gameStartTime;

    public abstract string MinigameID { get; }

    // Game Explanation Text is displayed to user through minigame UI
    public abstract string GameExplanationText { get; }

    public virtual bool HasGameBeenPlayedAndCompleted()
    {
        return gameHasBeenPlayed;
    }

    // StartGame() holds basic functionality that will happen on minigame
        startup, such as logging or setting start time
    public virtual void StartGame(){...}

    // GameOver() holds basic functionality that will happen on minigame
        end, such as completing logs and opening game menu
    public virtual void GameOver() {...}

    // VerifyLoggers will be implemented in each minigamemanager to verify
        that each metric logger has successfully loaded in the game scene
    protected abstract void VerifyLoggers();
}
```

Listing 3.3: Parts of the MiniGameManager.cs class

Platform Minigame

Platform Minigame places the player on a platform, thereby the name. Objects spawn at random intervals and move towards the player. The goal for the player is to avoid all red objects and hit all the green objects. No condition makes the player lose the game, but the goal is to get as high score as possible. The motivation behind not having a losing condition in this game was that it could be a very casual game. Thus, it will not hinder players from continuing gameplay even if they would perform a bit worse. This could help maintain flow in gameplay.

Game mechanics:

- Green and red objects spawn and move towards the player.
- Dodge red objects to not lose points.
- Hit green objects to score points.
- Objects can be hit with hands and head.

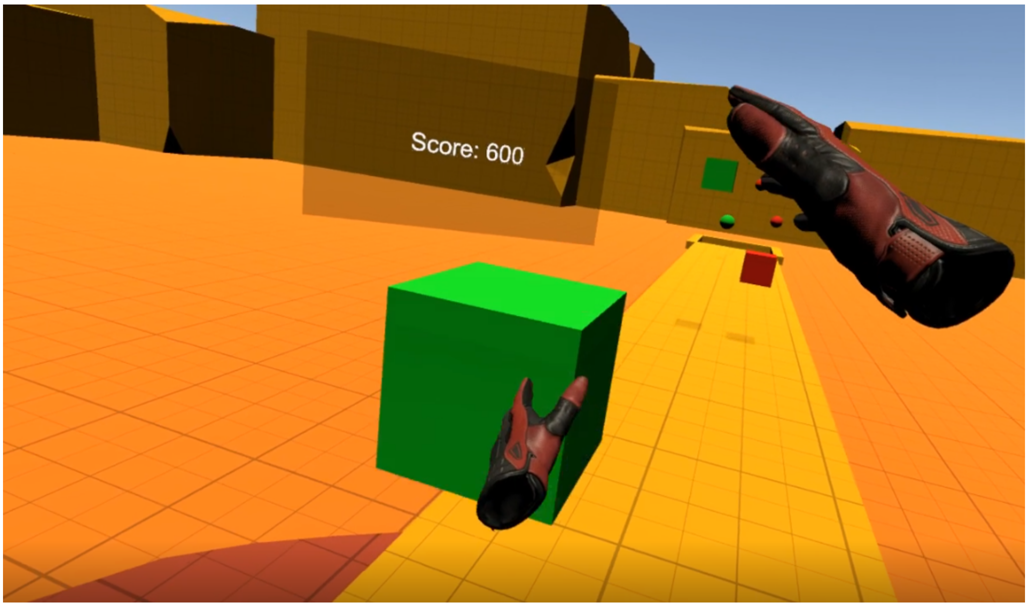


Figure 21: Screenshot from Platform Minigame

Drone Shooter

Drone shooter places the player in an urban sci-fi city environment. The player has five lives and is armed with a laser pistol. Drones spawn in waves and shoot missiles that have to be dodged by the player to not lose lives. The game is endless and has scalable difficulty with more difficult waves spawning over time.

Game mechanics:

- Drones spawn in waves. These shoot missiles at the player.
- Dodge missiles to not lose lives.
- Shoot drones with laser pistol to score points and progress in the minigame.

Figure 22: Screenshot from Drone Shooter Minigame

Reaction Time Trainer

The Reaction Time Trainer minigame is quite self-explanatory in its name. The game is based a lot on reaction time, with some hand-eye coordination and physical activity added to it. The game idea took inspiration from BATAK reaction time trainer instrument [11], a physical device that is for example used by football keepers to train their reaction time.

Game mechanics:

- 8 orbs float in front of the player, whenever one of these *enable*; they turn green and make a sound.
- When an orb is active; the orb must be hit quickly before it deactivates.
- If an inactive orb is hit, or an orb deactivates before hit; the player loses 1 life.
- Difficulty of the minigame will scale up over time, demanding quicker reactions from the player over time.
- When all 5 starting lives are lost, the game is over.

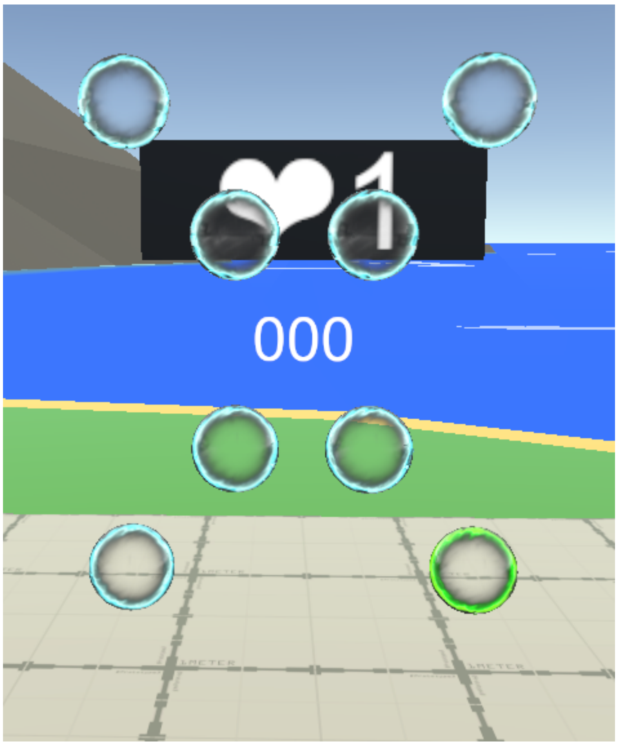


Figure 23: Reaction Time Trainer Minigame Overview - The bottom right orb is activated in this picture, and the player needs to hit it before it deactivates

[11]Youtube video with BATAK Reaction Test machine: youtube.com/watch?v=cyy4lqRERJM

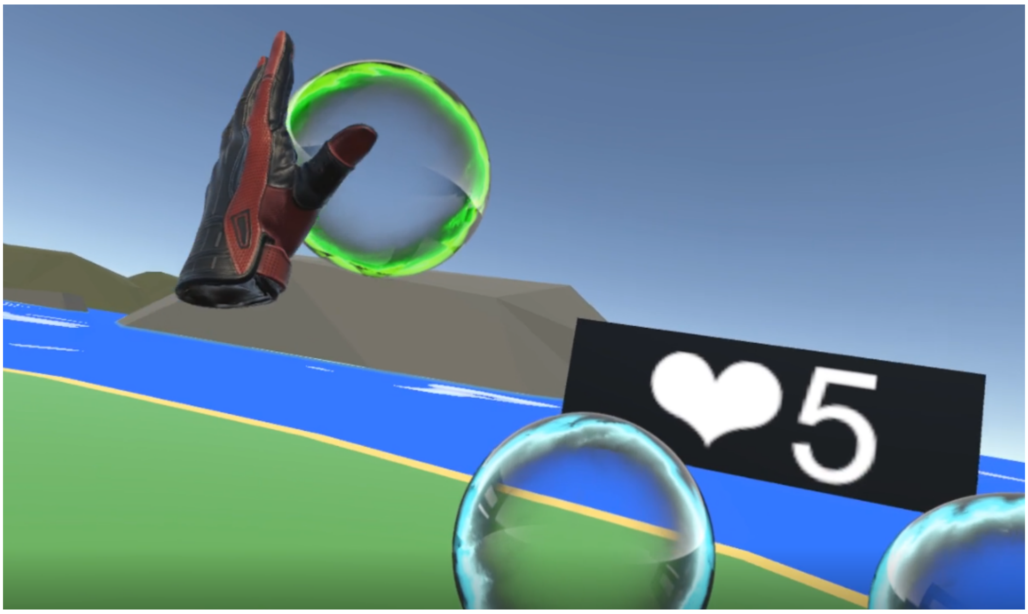


Figure 24: Screenshot from Reaction Time Trainer Minigame - Player reaching out for an active orb

3.4.4 Gamifications

The idea and motivations behind the gamifications added to the VR hub and minigames were to incentivize physical activity. In regards to serious games, several game elements can be added to help maintain flow and engagement in users.

> *Flow is a subjective state that people report when they are completely involved in something to the point of forgetting time, fatigue, and everything else but the activity itself.*
>
> - *Mihaly Csikszentmihalyi et al.* [62]

Music & Sound Effects

Every minigame had an independent list of songs that could be played during gameplay. When the player starts a minigame, a random song from the list is picked out and played. Each minigame has a separate list of songs, as the songs try to capture the atmosphere of the minigame as good as possible.

Additionally, the different objectives or interaction made by the player in minigames would have distinct sound effects to them, with the idea of making the interactions feel better for the player.

Haptic Feedback

Although it is a minor part of the feedback from minigames, I think that VR games benefit really well from immersion, and haptic feedback in controllers help facilitate this feeling. Any physical interaction with a game mechanic gives the user haptic feedback in the controller they interacted with. This meant that shooting the laser gun in *Drone Shooter*, or hitting a water orb in *Reaction Time Trainer* would hopefully feel better and more immersive to the player.

Achievements

Different players have different goals when playing games, and one important thing to consider when designing games is player types. A quite well-known classification of player types is *bartle's taxonomy* [63].

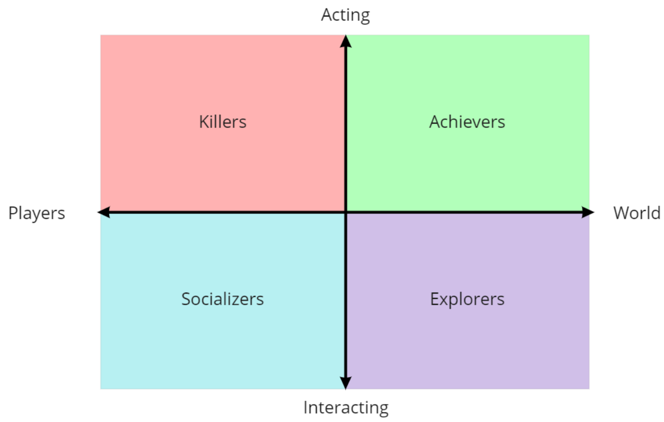


Figure 25: Bartle's Taxonomy (Player Types) [63]

To add an additional layer that caters to the *achiever* player type, it is a nice addition to have concrete achievements which the player can work towards. Most players will likely have some sort of feeling of reward when unlocking achievements, and not only the people categorizing as an *achievers*. I must acknowledge that the game is limited in its design for the *killer* and *socializer* player type though, as the game is designed as a single-player experience. Concerning the *explorer* player type, the main gamification for this player type is the island environment of the main hub. It allows players to walk around and explore this island, although it is somewhat limited in size and possible interactions. Figure 26 displays the different achievements that could be completed in the VR application.

Figure 26: Achievements seen from the main game menu

Difficulty Levels

All minigames had some sort of adjustable difficulty level. I found this as a significant addition because players need to have task difficulty that matches their skill level. If this was not present in the system, players could potentially feel bored if the gameplay was too easy, or anxious about playing if it was seen as too hard (as described by the flow model [64]).

Figure 27: Difficulty Selection in Minigame Menu

As fig. 27 displays, the player would be presented with *four* different difficulty levels to choose from. Additionally, the difficulty in *Reaction Time Trainer* and *Drone Shooter* minigames were scaling over time, making sure that there would always be a challenge for the player as the player would get more into the activity or in a state of *flow*.

Randomness in Minigames

All minigames had randomly generated content. This made each play-through of a minigame feel unique, and the idea was to increase the replayability of the different minigames.

It served its purpose in creating new experiences every time a minigame was played. However, some parts of the games could have benefited from having some scripted or predefined behaviour as it could add more flavour to the gameplay.

Listing 3.4 displays how one of the random functionalities for generating content worked. In this specific example, it is from the platform minigame, and this function displays how *hittable* or *obstacle* objects is added in a *grid* before moving towards the player.

```
public void AddObjects(System.Random rand, List<GameObject> hittables,
    List<GameObject> obstacles, int difficultyLevel)
{
    // Fuzzy logic for amount of objects to spawn
    // Minimum 1, max 2 at easy + medium difficulty
    // Minimum 1, max 3 at hard + extreme difficulty
    int objectsToAdd = 1 + (rand.Next() % ((2 + (difficultyLevel / 3))));

    for (int i = 0; i < objectsToAdd; i++)
    {
        GameObject objectToSpawn;
        if (rand.Next() % 2 == 0) // 50/50 to spawn hittable or obstacle
        {
            objectToSpawn = hittables[rand.Next() % hittables.Count];
        }
        else
        {
            objectToSpawn = obstacles[rand.Next() % obstacles.Count];
        }

        // Find all nodes in grid that have no child objects
        List<Transform> availableNodes = childNodes.FindAll(x =>
            x.childCount == 0);

        // Instantiate object with one of these available parent nodes
        Instantiate(objectToSpawn, availableNodes[rand.Next() %
            availableNodes.Count]);
    }
}
```

Listing 3.4: AddObjects() function for randomly generating objects in Platform Minigame

3.4.5 Assets & Proprietary Assets

For the VR application, I have used a mix of personally made assets, in combination with proprietary assets. Most of the 3D models and sounds are third-party assets, as I am a software engineer and not a designer or sound producer. However, some 3D models have been created with ProBuilder[12] in Unity.

When it comes to programming software, it is often seen as unnecessary to "*reinvent the wheel*". Unless, the person is aimed at learning something new, or improving an existing solution. With the increased prevalence of open-source software recently, this has become even easier for developers all over the world.

However, the state of open-source for game programming is not as established yet. Many libraries and tools created for game engines are still proprietary assets that developers have to buy to get access to, or their tools are free but internally owned by a company. The third-party assets I used for the VR application is listed below.

List of third-party code assets:

- ProBuilder - Unity 3D modelling
- Easy Save 3 - Save system (Proprietary)[13]
- SineVFX - Particle & Shader Effects (Proprietary)[14]
- SteamVR - VR Integration[15]

[12] unity3d.com/unity/features/worldbuilding/probuilder
[13] assetstore.unity.com/packages/tools/input-management/easy-save-the-complete-save-data-serialization-asset-768
[14] assetstore.unity.com/packages/vfx/particles/spells/forcefield-effects-123431
[15] assetstore.unity.com/packages/tools/integration/steamvr-plugin-32647

3.4.6 Logging Game Sessions

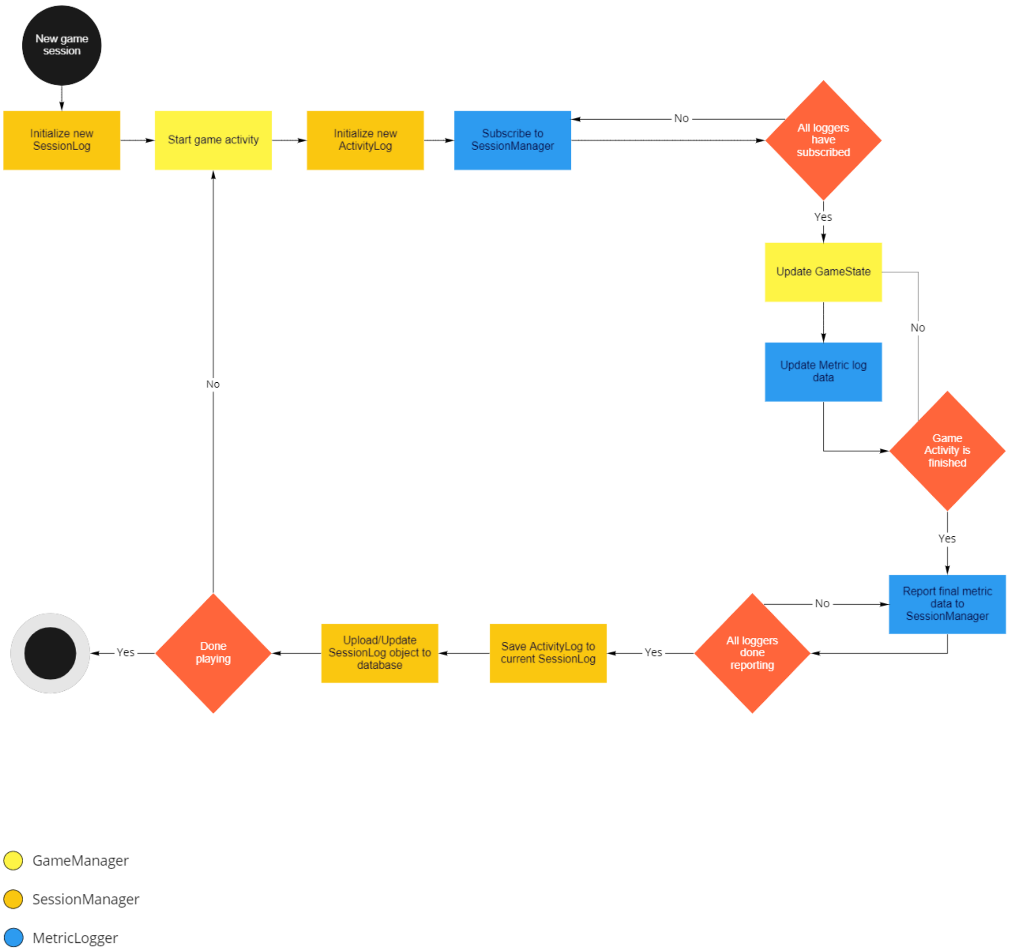


Figure 28: Logging of activities in the VR application

Figure 28 displays an activity diagram describing the process of logging game activities in the VR application. It is a continuous process that begins as soon as the player starts a game activity.

It starts with the `SessionManager` initializing a new `Session` when the application boots up, called from the `Awake()`-event in Unity's execution order[16]. The full code from the referenced code below can be seen in the VR application repository.[17]

[16] docs.unity3d.com/Manual/ExecutionOrder.html
[17] SessionManager script:
gitlab.com/akerholten/exerisland-vr-exergame-hub/-/blob/master/Assets/_Assets/Scripts/Monitoring/SessionManager.cs
MetricLogger script:
gitlab.com/akerholten/exerisland-vr-exergame-hub/-/blob/master/Assets/_Assets/Scripts/Monitoring/MetricLogger.cs

```
private Session currentSession;
private string sessionID = "";
private Activity currentActivity;

private void Awake()
{
    if (instance == null)
    {
        instance = this;
        InitNewSession();
    }
    else if (instance != this)
    {
        Destroy(gameObject);
    }
    DontDestroyOnLoad(this); // We want the singleton object to persist
        through scenes
}

public void InitNewSession()
{
    currentSession = new Session();
    currentActivity = null;
}
```

Listing 3.5: SessionManager instantiating a session log in Awake() call from Unity

On game activity start, the game manager calls the `SessionManager`'s `StartNewActivity()` function. This creates a new `Activity` object that is stored in the `SessionManager` while *metrics* are logged during game play.

```
public void StartNewActivity(string minigameID)
{
    if (hasActivity) // if we previously had an activity we have not
        uploaded yet
    {
        Debug.LogWarning("An activity was not dealt with before a new one
            started");
        ActivityCompleted();
    }
    currentActivity = new Activity(minigameID);
    hasActivity = true;
    sessionDataUploaded = false;
    metricLoggers = new List<MetricLogger>();

    // Calling that the event has started
    ActivityStarted(this, new ActivityArgs(minigameID));
}
```

Listing 3.6: SessionManager - StartNewActivity()

When the `ActivityStarted()`-event is invoked, all `MetricLoggers` will subscribe to the `SessionManager` such that their data is linked to the current activity when completed.

```
public abstract class MetricLogger : MonoBehaviour
{
    protected void Awake()
    {
        // Subscribe to ActivityStarted Event
        SessionManager.ActivityStarted -= OnActivityStarted;
        SessionManager.ActivityStarted += OnActivityStarted;
    }

    protected void OnActivityStarted(object sender, ActivityArgs e)
    {
        currentMetric = InitiateMetric();
        SessionManager.instance.SubscribeLogger(this);
        active = true;
        OnLoggingStarted();
    }

    protected abstract void OnLoggingStarted();
}
```

Listing 3.7: MetricLogger - OnActivityStarted()

When the game activity is finished, the `SessionManager` goes through each logger and add their *metrics* to the current `Activity`-object, before uploading it.

```
public void ActivityCompleted()
{
    if (currentActivity != null)
    {
        foreach (var logger in metricLoggers)
        {
            AddMetricToCurrentActivity(logger.ReportMetricAndStopLogging());
        }
        currentSession.activities.Add(currentActivity);
    }

    currentActivity = null;
    hasActivity = false;

    VerifySessionAndUpload();
}
```

Listing 3.8: SessionManager - ActivityCompleted()

Lastly, the session is sent to the backend for storing in the database. This is done through a HTTP *POST*-request as seen in listing 3.9. In addition to the session object itself, it contains the user's personal ID as authorization to know which user the data should be stored to.

```
public void VerifySessionAndUpload()
{
    if (HasDataToUpload && !sessionDataUploaded && !currentlyUploading)
    {
        StartCoroutine(UploadSessionToDB());
    }
}

public IEnumerator UploadSessionToDB()
{
    currentlyUploading = true;

    // For updating the top-level "duration" value in json object before
        sent out
    currentSession.UpdateDuration();

    string jsonRequest = JsonUtility.ToJson(currentSession);

    byte[] jsonReqData = System.Text.Encoding.UTF8.GetBytes(jsonRequest);

    UnityWebRequest req = new UnityWebRequest();

    req.method = "POST";
    req.uploadHandler = new UploadHandlerRaw(jsonReqData);
    req.downloadHandler = new DownloadHandlerBuffer();

    req.SetRequestHeader("Personal-ID", GameManager.instance.UserID);
    req.SetRequestHeader("Content-Type", "application/json;
        charset=utf-8");

    // If the session has not been initialized / uploaded before
    if (sessionID == "")
    {
        req.url = Constants.BACKEND_URL + Constants.Paths.UploadSession;
    }
    else // Session has been uploaded before, so we just update the
        existing session
    {
        req.url = Constants.BACKEND_URL +
            Constants.Paths.UpdateExistingSession;
        req.SetRequestHeader("Session-ID", sessionID);
    }

    yield return req.SendWebRequest();

    if (req.isNetworkError)
    {
        Debug.Log("Error while sending session: " + req.error);
    }
    else
    {
        // If sessionID is not stored from backend yet, we do it now,
            basically means that it is the first push of the session
        if (sessionID == "")
        {
            sessionID = req.downloadHandler.text;
        }
        sessionDataUploaded = true;
    }
    currentlyUploading = false;
}
```

Listing 3.9: SessionManager - Upload session functionality

3.5 ExerIsland: Web Frontend for Activity Tracking

The frontend application[18] developed for web with Flutter contains three main screens; observer dashboard, participant dashboard and a session view screen. These screens and their purposes is explained below in sections 3.5.2 to 3.5.4. In addition to these main screens, the app contains a signup-screen, login-screen and an error-screen.

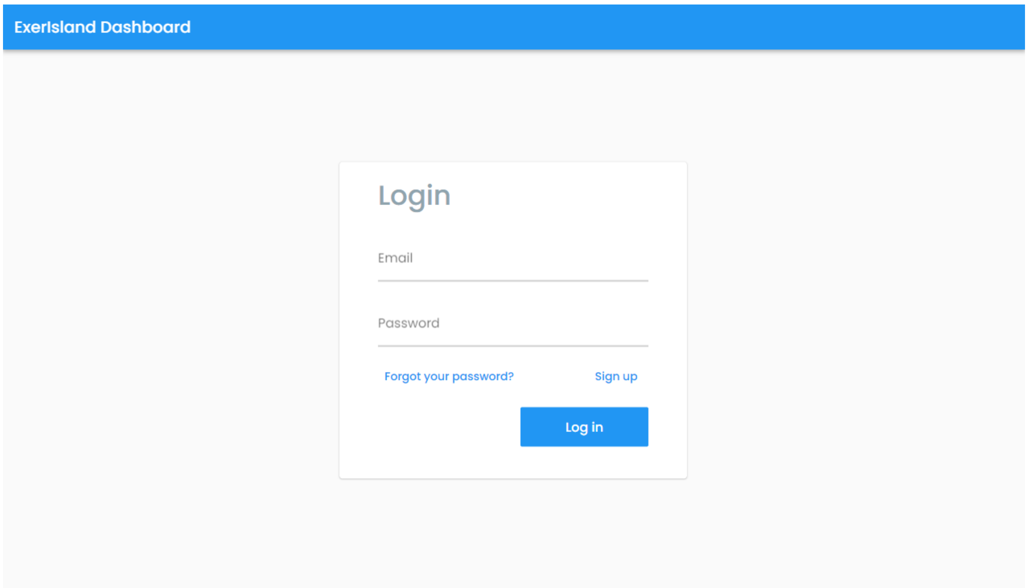


Figure 29: Frontend: Login Screen

3.5.1 UI & Widgets in Flutter

To give the reader a quick insight into how User Interface (UI) in Flutter works, this is briefly described here.

UI in Flutter is built up through *widgets*, inspired by how *React*-applications are implemented [65]. The main idea is that all UI is built up through a widget-tree. This widget-tree describes what the view will look like given its configuration, state and context. Whenever the *state* changes, the widget rebuilds with the new state.

Listing 3.10 displays a minimal example of a Flutter app in Dart taken from the flutter documentation [65]. The `runApp()`-function takes the given widget and makes it the root of the widget tree. In the example here, there are two widgets, the `Center`-widget and its child object, the `Text`-widget. This hierarchy of widgets is what builds up the rendered UI in Flutter applications.

[18]The deployed frontend application for ExerIsland can be reached at exerisland.com

```dart
import 'package:flutter/material.dart';

void main() {
  runApp(
    Center(
      child: Text(
        'Hello, world!',
        textDirection: TextDirection.ltr,
      ),
    ),
  );
}
```

Listing 3.10: Hello World Example App in Flutter

With widgets, one can quickly create many reusable UI components, enabling great speed of rapid prototyping within the Flutter framework. This was one of the key reasons for choosing Flutter for this frontend application. However, this UI-building with widgets tends to create very cluttered and large trees of components. This can make the code less readable and harder to interpret unless one understands the underlying functionalities of the framework. I often see this as an issue for any web development framework. There are also similar issues with javascript or typescript in *React* for example.

In my frontend application, I acknowledge that there could likely be improvements to make for readability. Yet, I would like to say that web development at its core tends to not make this process that easy. Examples of built up *widgets* for my web application can be seen in full code appendix B.3.

3.5.2 Observer Dashboard

The observer dashboard is the dashboard used by *researchers* in the system. As seen in fig. 30 it displays a list of all participants the observer has access to, and it also displays some meta-information about these users.

For the observer's perspective in the dashboard, the idea is that it gives an overview of how study participants are progressing along. The observer can check in on the web dashboard to see whether participants seem to struggle with a specific task, or get some early insights into the data generated by participants in the study.

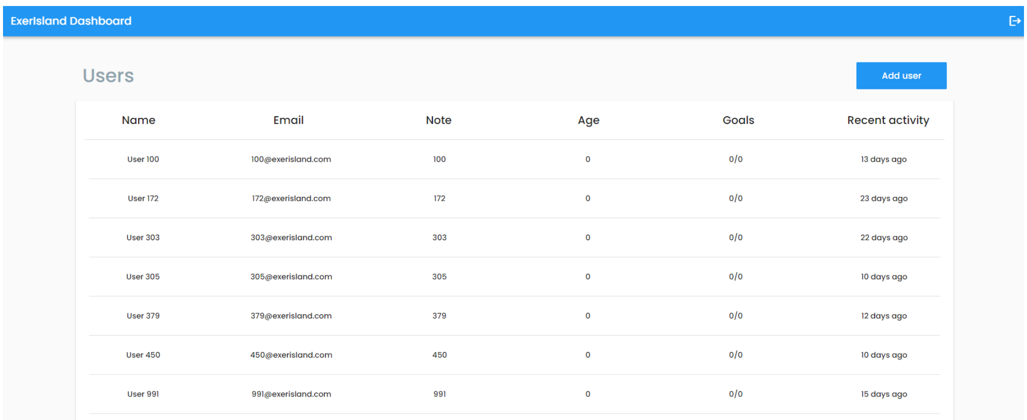


Figure 30: Frontend: Observer Dashboard

By clicking on an individual participant from the list in the dashboard, the observer will be presented with the participant's dashboard.

3.5.3 Participant Dashboard

The participant dashboard is the main entry point for participants logging in to the web application. An observer with access to the specific participant can also see this dashboard.

From the participants' perspective, the main idea of this dashboard is to be able to analyze and look at personal statistics in retrospect. The idea was that this could potentially motivate participants to further engage in physical game activities in VR. This was directly relevant to*research question 2: "How can exercise trackers motivate people to engage more in physical activity?"* and *research question 3: "How should data from games be presented to engage users more in activities?"*

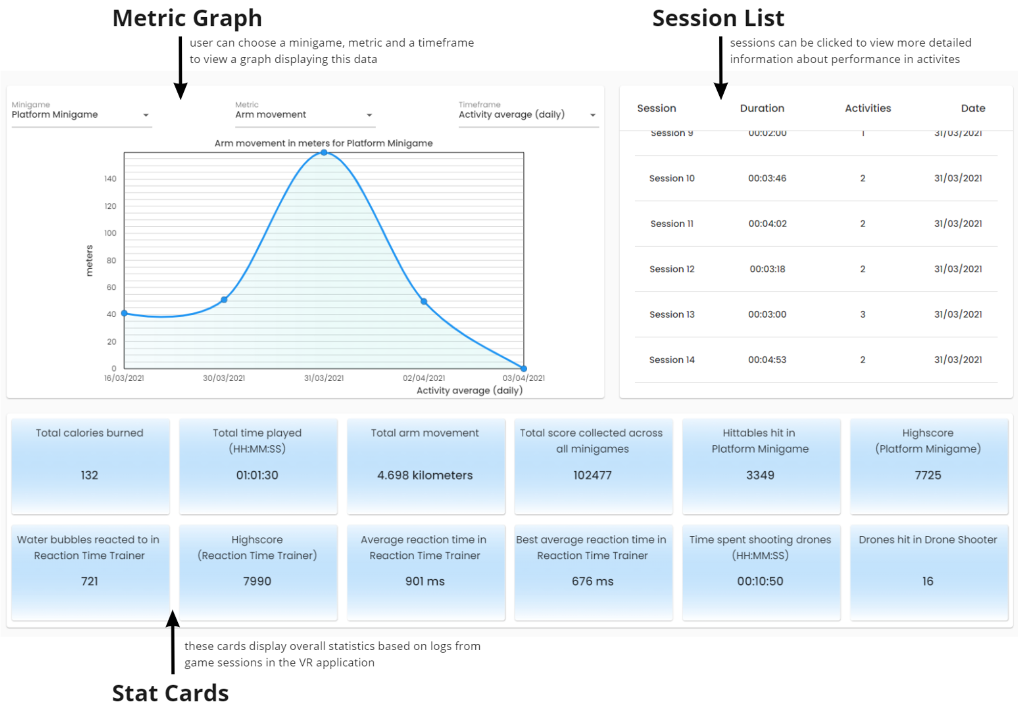


Figure 31: Frontend: Participant Dashboard

Figure 31 displays a screenshot of the participant dashboard. There are three main components displayed in this dashboard; metric graphs, session list, and statistic cards.

Metric Graphs

The user can select a minigame, a metric from the chosen minigame, and a timeframe accordingly. The web application will then generate and display a graph with this data.

Session List

The session list displays all sessions the user has logged from the VR application since the beginning of the study. Clicking an item in this list will send the user to the session view of the specific session (see section 3.5.4 below).

Statistic Cards

Some general statistics that are collected from the data is displayed through these statistic cards in the participant dashboard. Some data are only from specific minigames, while other data is joint from all different types of activities the participant has performed.

3.5.4 Session View

Figure 32 displays how the session view looks like. Each activity the user has performed can be seen on the left side. By clicking on any of these activities, the metric data displayed on the right will update with the ones from the specific activity selected. This allows users to inspect all metrics logged from all activities during a single session.

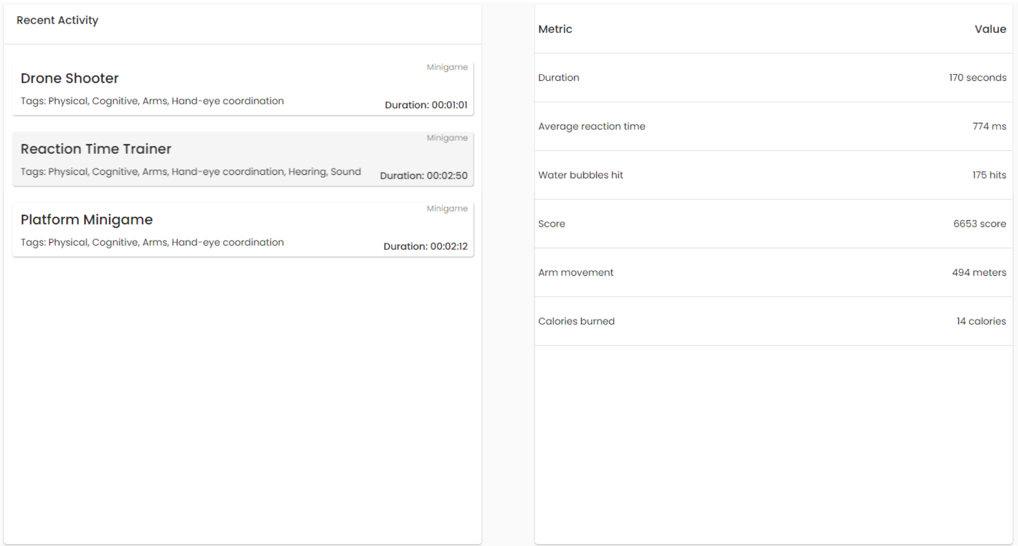


Figure 32: Frontend: Session View

3.6 ExerIsland: Enabling Remote Study Experiments

To successfully create a system for remote study experiments, there were several design and implementation decisions made in the development process. In this section, certain features that were required to enable the ability to run remote user studies are explained.

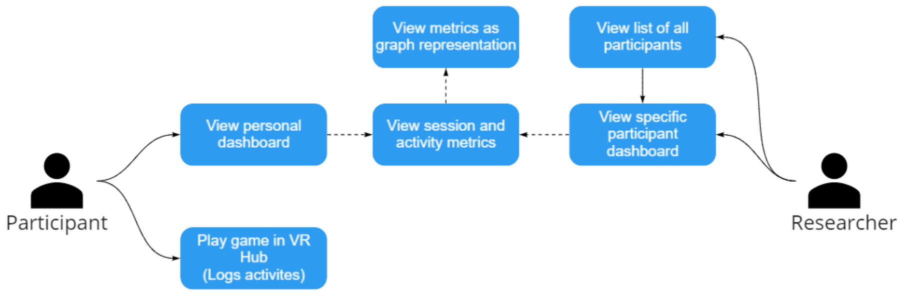


Figure 33: High-Level Use-Case diagram for researcher and participant

Figure 33 displays the high-level use-cases that exist in the ExerIsland platform. The initial idea for the platform was designed around the relationship between a healthcare professional and a patient undergoing rehabilitation. However, I discovered quite early that this relation was easily transferable to a *researcher - participant* perspective.

In its essence, the researcher can use the platform to view the data from all people participating in a study. The participant, on the other hand, can log data through playing games in the VR application, and view their personal dashboard.

I originally intended to have a *task*-module for the ExerIsland platform as well. This would be the list of tasks the participants were expected to complete during the study and could be retrieved from the participant within VR. This could potentially have made it even easier for researchers to keep track of progress from participants. It would also make it easier for participants to keep track of their current tasks to perform for the study.

However, due to time limitations for development, I decided to discard this functionality as it would be more or less a Quality-of-Life (QoL) feature. Instead, the tasks in the user study were distributed through task sheets (described in section 4.2.3).

3.6.1 Non-identifiable Participant ID

One important thing to consider for research conducted with people is; how can you maintain anonymous participation? In the case of the ExerIsland platform, this was also an important consideration to make. There had to be a system in place which made it so that people could partake in a study without having their data compromised of any sort. To understand the incentives better, I have listed a few conditions that had to be met:

1. All data from a participant, from any source in the study, must be possible to link together for data analysis.
2. Data from a participant must never be linked to any personally identifiable data.
3. The researcher must, *under no circumstance*, be able to identify a participant through their data.

Essentially, there had to be a way to generate a personal ID that would link to data logs from the VR application, and which the subject could use when answering research surveys. With that system in place, all data from the participant could always be linked together. Additionally, as long as this ID was never exposed by the participant, their *anonymous participation* would be maintained.

User Generation

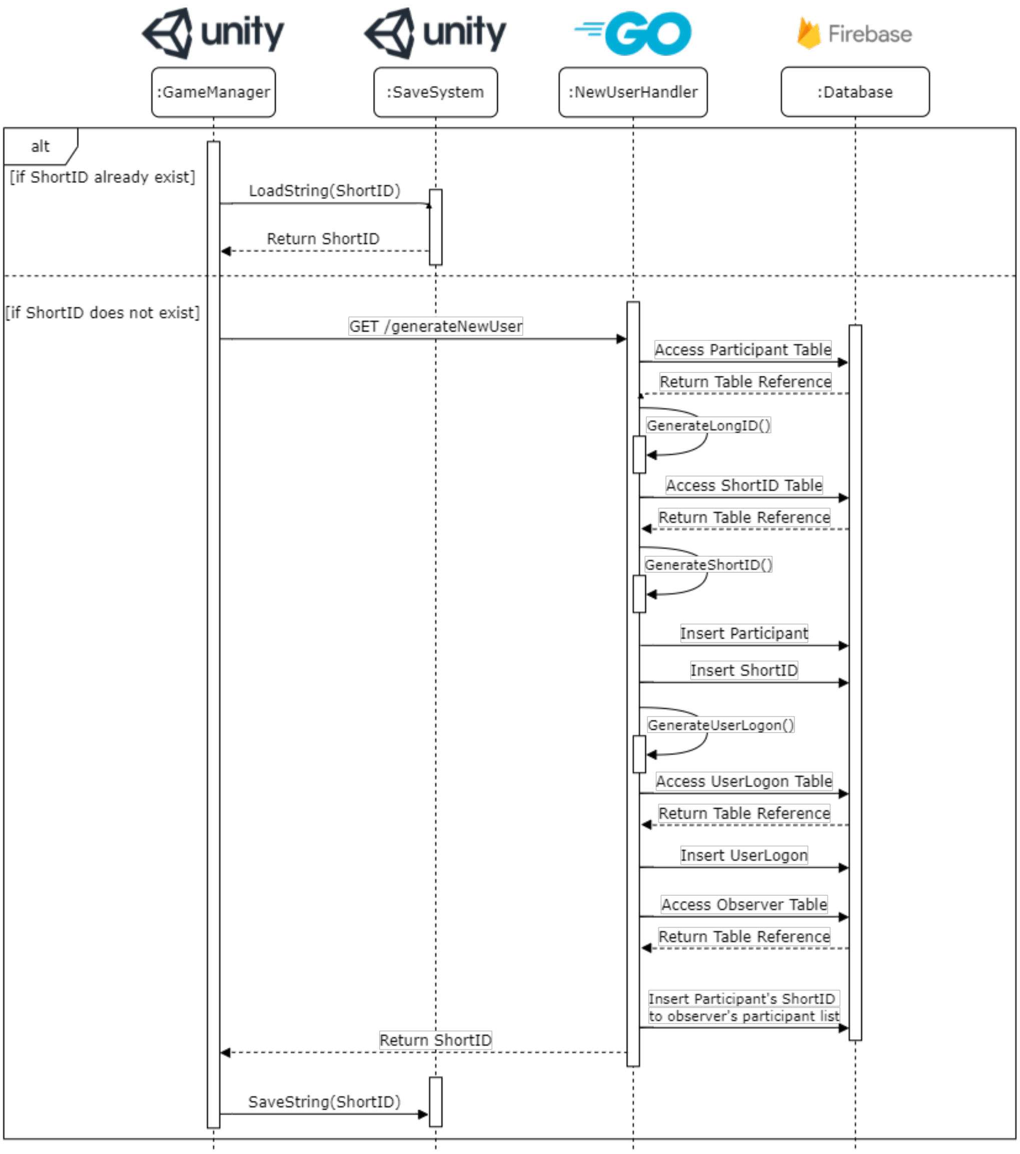


Figure 34: Sequence diagram displaying how a new participant user is generated in the ExerIsland platform

Figure 34 displays the process of generating a new participant user in the system. It starts when the VR application boots up. The `GameManager` singleton calls their `public void CollectUserID()` function. This function checks whether a *UserID* for the participant already is saved or not.[19]

```
// UserID is displayed in game UI (table in main hub)
public string UserID => _userID;
private string _userID = "";

public void CollectUserID()
{
    if (ES3.KeyExists(USERID_SAVE))
    {
        _userID = ES3.LoadString(USERID_SAVE, "UserID loading...");
        if (_userID == "UserID loading..." || _userID == "")
        {
            Debug.LogWarning("Something went wrong loading user ID, we try
                create a new user");
            StartCoroutine(CreateNewUser());
        }
    }
    else
    {
        // We need to call the backend to create a user since it does not
            exist
        StartCoroutine(CreateNewUser());
    }
}
```

Listing 3.11: GameManager.cs - CollectUserID()

If a *UserID* is not stored in the application already, the `CreateNewUser()`-coroutine is started. The coroutine will send the *GET /generateNewUser* web request to the backend and retrieve the *ShortID* generated on the backend. The ID is stored in the VR application's persistent storage, for logging future activities in the application.

```
private IEnumerator CreateNewUser()
        {
            UnityWebRequest req = new
                UnityWebRequest(Constants.BACKEND_URL +
                Constants.Paths.GenerateNewUser);
            req.method = "GET";
            req.downloadHandler = new DownloadHandlerBuffer();

            yield return req.SendWebRequest();

            if (req.isNetworkError)
            {
                Debug.Log("Error while sending session: " + req.error);
                _userID = "UserID load error";
            }
            else
            {
                Debug.Log("Received when uploading session: " +
                    req.downloadHandler.text);
                _userID = req.downloadHandler.text;
                ES3.Save(USERID_SAVE, _userID);
            }
        }
```

Listing 3.12: GameManager.cs - CreateNewUser()

[19]Full code examples from the referenced code here can be seen here:
gitlab.com/akerholten/exerisland-vr-exergame-hub/-/blob/master/Assets/_Assets/Scripts/Managers/GameManager.cs
gitlab.com/akerholten/vr-health-and-wellness-remote-monitoring/-/blob/master/backend/internal/tools/tools.go

On the backend, a new participant user is initiated. Their identification throughout the user study will be the *ShortID* created through a simple random-generation implementation in Golang:

```
var (
        base62Chars string =
            "0123456789ABCDEFGHIJKLMNOPQRSTUVWXYZabcdefghijklmnopqrstuvwxyz"
)

func GetNewShortUniqueID(length int, safetyCount int64) string {
        newId := ""

        seed := time.Now().Unix() + safetyCount

        rand.Seed(seed)
        for i := 0; i < length; i++ {
                newId = newId + string([]rune(base62Chars)[rand.Int()%10])
                    // 36 for only upper-case // % 10 for only numbers
        }

        return newId
}
```

Listing 3.13: tools.go - GetNewShortUniqueID(), inspired by implementations on Stack Overflow [66]

The `safetyCount int64` from listing 3.13 is a safety *count* mechanism, used to ensure that several seed generations can take place in case the returned ID already exists in the database. While the `length int` variable tells how long the ID generated should be. To make it simpler for participants to recall their ID, I used a length of *3* for the ID in the study, and with only digits as the possible characters.

The backend also makes a call to generate a *Logon_User* such that these participants can log in and view their personal dashboard in the web application. For simplicity, the participants could log in with their *ShortID* as username and password. I realize that this is a substantial security flaw of the system, but network security was never the main focus area of this thesis project. Thus, it made the platform partly limited in certain security features, as it was more important to focus on getting the functionality of remote user studies implemented. The issues around the ID-system and potential solutions are discussed in section 6.1.1.

After the participant user has been generated and stored in the database, the backend will also access the administrator *observer* stored in the database. The new participant's *ShortID* is then stored into the observer's list of participants, providing them access to view this participant in their dashboard.

Handing ID to the User

To distribute the *participant-ID* to the user, this is done within the VR application. The participant is presented with a wooden board that displays their personal ID as seen in fig. 35 below. The information board also tells the user explicitly to keep this identity private and write it down for themselves. That way, they can easily access it when required for other parts of the user study (e.g. when answering questionnaires).

Figure 35: Displaying participant ID to the subject in VR

3.6.2 Software Security Aspects

I have limited myself in the amount of time I can spend on security features, to not deviate too much from the main focus area. The *ExerIsland* platform was mainly developed for remote user study purposes after all. Nevertheless, there are a few security features implemented, and some of these are described here. [20]

Salting & Hashing

When a new user is created, their password is salted and hashed before storing it in the database. We can take a look at the `SignupHandler` dealing with a new signup-request below in listing 3.14.

[20] Full code examples from the referenced code here can be seen here:
gitlab.com/akerholten/vr-health-and-wellness-remote-monitoring/-/blob/master/backend/internal/tools/tools.go
gitlab.com/akerholten/vr-health-and-wellness-remote-monitoring/-/blob/master/backend/internal/db/authenticationInterface.go
gitlab.com/akerholten/vr-health-and-wellness-remote-monitoring/-/blob/master/backend/internal/handlers/observerHandlers.go

```
func SignupHandler(w http.ResponseWriter, r *http.Request) {
	log.Printf("Got a signup request...")
	defer r.Body.Close()
	if r.Header.Get("Content-Type") != "application/json;
	    charset=utf-8" {
		w.WriteHeader(http.StatusBadRequest)
		return
	}

	var signupData db.SignupUser

	respBody, err := ioutil.ReadAll(r.Body)
	if err != nil {
		w.WriteHeader(http.StatusForbidden)
		return
	}

	err = json.Unmarshal(respBody, &signupData)
	if err != nil {
		w.WriteHeader(http.StatusBadRequest)
		log.Printf("Error unmarshaling: %s, error: %v",
		    string(respBody), err)
	}

	// Data input validation happens here
	// ...
	// ...

	valid, err := validator.ValidateStruct(signupData)
	if err != nil || !valid {
		w.WriteHeader(http.StatusInternalServerError)
		return
	}

	ctx := context.Background()

	exists, err := db.IsExistingUser(signupData.Email, ctx)
	if err != nil {
		log.Panicf("Error: %v", err) // panic returns out
	}
	if exists { // If the user actually already exists, it can't be
	    created again
		w.WriteHeader(http.StatusConflict)
		w.Write([]byte("Account with that email already exist"))
		return
	}

	// Salting and hashing password
	signupData.Password = tools.ConvertPlainPassword(signupData.Email,
	    signupData.Password)

	if signupData.UserType == constants.ObserverType {
		// Adding to observer table for the observer data
		longId, err := db.AddToObserverTable(signupData, ctx)
		if err != nil {
			log.Panicf("Error: %v", err)
		}
		signupData.UserID = longId

		// Adding to User_Logon table for login
		err = db.AddToUserTable(signupData, ctx)
		if err != nil {
			log.Panicf("\nUser was not added to user table,
			    error: %v\n", err)
		}
	}

	w.Write([]byte("User added!"))
}
```

Listing 3.14: authenticationHandlers.go - SignupHandler()

The ConvertPlainPassword()-method (see listing 3.15 below) takes in the email and password. The email is used as the *salt* in this case, for the sake of having a simple way to do salting. The email is hashed before being input as salt to the raw password. The outcome of this is hashed once more before stored in the database. The hash method used here is *SHA-512*, which can be considered unnecessary, but was done to minimize any risk of hash collisions in passwords.

```
// ConvertPlainPassword hashes a raw password and returns the hashed
    password
func ConvertPlainPassword(rawEmail, rawPassword string) string {
        hashedEmail := CreateHash(rawEmail)
        return CreateHash(hashedEmail + rawPassword)
}

// CreateHash creates a new hash string
func CreateHash(key string) string {
        hasher := sha512.New()
        _, err := hasher.Write([]byte(key))
        if err != nil {
                log.Panicf("Error writing hash, %v", err)
        }
        return hex.EncodeToString(hasher.Sum(nil))
}
```

Listing 3.15: tools.go - Salting and Hashing Passwords

Authentication & Authorization

For logging in to the platform, the user is authenticated directly by looking up the sent data in the database as seen in AuthenticateUser()-method below. Their password is first salted and hashed using the similar procedure as described above for user creation.

```
// AuthenticateUser verifies that the user exists in the database
    User_Logon entry with email and password. Returns userID or error
func AuthenticateUser(form LoginForm, ctx context.Context) (*User_Logon,
    string, error) {
        table := DBClient().Database.NewRef(TableUser)

        results, err := table.OrderByKey().GetOrdered(ctx)
        if err != nil {
                return nil, "", err
        }

        for _, r := range results {
                var userEntry User_Logon
                if err := r.Unmarshal(&userEntry); err != nil {
                        return nil, "", err
                }

                // If there is a match entry, authentication is successfull
                if form.Email == userEntry.Email && form.Password ==
                    userEntry.Password {
                        return &userEntry, r.Key(), nil
                }
        }

        return nil, "", nil
}
```

Listing 3.16: authenticationInterface.go - AuthenticateUser()

For *authorization*, this is mainly done through using the *cookie* stored in the client's session. We can take a look at an example function that uses authorization below:

```
func GetParticipantsHandler(w http.ResponseWriter, r *http.Request) {
        log.Printf("Got a request for all participants...")
        defer r.Body.Close()

        ctx := context.Background()

        // Fetch cookie which will be used for authorization
        clientCookie, err := cookie.FetchCookie(r)
        if err != nil {
                // This could mean that the cookie is not present so
                    technically not a internal server error, but could be
                    bad request
                log.Printf("Could not fetch cookie, err was: %v", err)
                w.WriteHeader(http.StatusBadRequest)
                return
        }

        // Authorize user with their fetched cookie
        user, err := db.GetUserFromCookie(clientCookie, ctx)
        if err != nil {
                log.Printf("Could not fetch user from cookie, err was:
                    %v", err)
                w.WriteHeader(http.StatusBadRequest)
                return
        }
        if user == nil {
                log.Printf("Could not fetch user from cookie, err was:
                    %v", err)
                w.WriteHeader(http.StatusBadRequest)
                return
        }
        if user.UserType != constants.ObserverType {
                w.WriteHeader(http.StatusUnauthorized)
                return
        }

        // Get array of participants from observerInterface function
        participants, err := db.GetParticipants(clientCookie, ctx)
        if err != nil {
                log.Printf("Could not fetch participants from this user,
                    err was: %v", err)
                w.WriteHeader(http.StatusInternalServerError)
                return
        }

        // Marshal it into json and return
        participantsJson, err := json.Marshal(participants)
        if err != nil {
                log.Printf("Could not marshal participants from this user
                    before return, err was: %v", err)
                w.WriteHeader(http.StatusInternalServerError)
                return
        }

        w.WriteHeader(http.StatusOK)
        w.Write(participantsJson)
}
```

Listing 3.17: observerHandlers.go - GetParticipantsHandler()

The `GetParticipantsHandler` first fetches the cookie from the client and decodes it. Further, the data from the cookie is used to try to fetch a user from the database. If it is successful at retrieving a user from the cookie, it then checks that the user fetched is a *observer*-user in this scenario. This is because the `GetParticipantsHandler` is only supposed to be accessible for observers, and is only usable for observers also since only they can have a list of participants connected to them.

If everything is successful, the backend calls for a lookup in the database to access the participants of the user. This result is marshalled into a JSON object before written into the HTTP response.

These examples display simple but effective tools for having some layers of security within the platform. Though there was limited focus on software security, I feel that for the use-case of the platform, it was secure enough.

4 Experiment Methodology

In this thesis, there were two study experiments conducted. The first part is a questionnaire study collecting valuable data for understanding people's tendencies and experiences with exercise trackers. This is relevant information for *focus B: Gamification techniques for exercise motivation*.

The second experiment was a two-week user study conducted with the developed ExerIsland platform. The experiment results from this study are also relevant to *focus B*. Furthermore, the evaluation of the platforms ability to enable these remote user studies is most applicable to *focus A: Technologies to enable remote user studies*.

4.1 Questionnaire Study on Exercise Trackers

The data from the two-week experiment with the *ExerIsland* platform could give appealing findings related to exercise motivation and the potential effects of the personal tracking dashboard. Nevertheless, I estimated that it would be a hard task to get many participants in such a comprehensive study, which also required users to have access to VR equipment.

In other words; I wanted to have a secondary data collection experiment, that could bring me a lot more data to analyze in regards to *focus B*. This was the main motivation for creating and distributing this online questionnaire study. The benefits are that it is simple to distribute, and it can also be easier to get a more wide demographical sample with an online study.

A drawback of an online questionnaire is that participants have a harder time asking questions. There could for example be some questions or information in the survey they do not understand. Additionally, the results of the survey are always based on the subjective perspective of the participant, meaning that there are several potential biases to consider when analyzing the data.

4.1.1 Questions

The questionnaire was created with Nettskjema[1]. All questionnaire questions can be seen in appendix C[2]. The time to complete the survey was estimated to be around 10-15 minutes, based on a *pilot*-test done with 9 people for the survey.
The main data collection points are divided into *4* categories:

Demographics

Information such as age and sex of the participant.

Relation to Exercise

Data about the participant's relation with physical activity and exercise. Questions are about exercise length, amount, and also preferences in activities.

[1] nettskjema.no
[2] Questionnaire questions can also be seen here: github.com/akerholten/MSc-Thesis/blob/main/OnlineQuestionnaire_ExerciseTrackers/OnlineQuestionnaire_ExerciseTrackers.pdf

Social Media & Videogame Relation

Questions regarding the person's relation to social media and videogames. This could potentially give interesting insights into whether this does affect how persuasive exercise trackers is at people with different tendencies with these media.

Experiences & Thoughts around Exercise Trackers

The main part of the questionnaire consists of questions regarding people's experiences and thoughts about exercise trackers.

Cross-Validation of Data

Some of the questions in the questionnaire are similar to each other. This is a decision by design. It is done to try to cross-validate the answers from participants.

An example of such cross-validation is with these two questions in figs. 36 and 37:

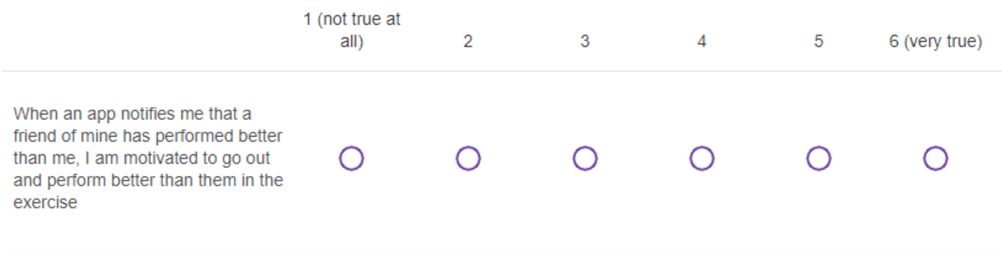


Figure 36: Likert scale from questionnaire about motivation by competition

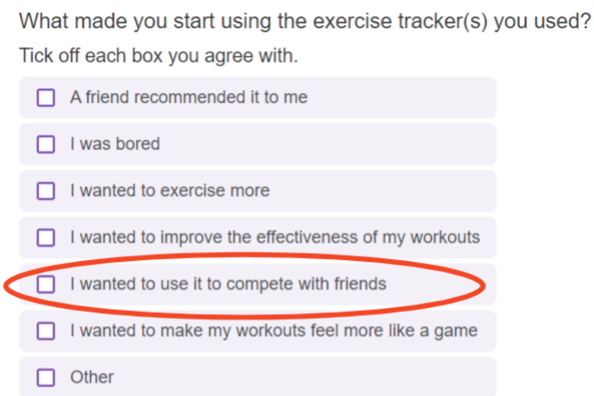


Figure 37: Check-box from questionnaire about motivation by competition

The questions are framed slightly different, but essentially they are asking about the same thing. Questions like this can further strengthen the value of the data as it works as cross-validation, to get a more ensured idea of what the participant's personal opinion is.

4.1.2 Querying for Participants

To query for participants to this online questionnaire, I distributed the questionnaire in my personal network and in channels related to NTNU and research.

I distributed the questionnaire on these platforms:

- Facebook [3]
- Reddit: *r/ntnu* [4] and *r/SampleSize* [5] subreddits.
- NTNU's Innsida [6] (for students)
- NTNU Gjøvik's Discord [7] channels

Through all of these channels, there was a total of *94* participants who responded to the survey. Subjects were aged 18-70 years old. There were 40 female, 53 male, and 1 respondent who did not indicate their *sex* in the demographics. 43 participants were aged between 18-24 years old.

4.1.3 Data Analysis

To analyze the results collected through the questionnaire, I used SPSS analytics software from IBM [8]. The results from this analysis is described in section 5.1.

4.2 Two-Week Experiment with ExerIsland Platform

The two-week user experiment with the developed ExerIsland platform set out to investigate a few different things. Concerning *focus A: Technologies to enable remote user studies*, this practical experiment would be highly relevant to evaluate whether the exergame platform was able to enable remote user studies. With that, the practical experiment would also shed light on what potential benefits and drawbacks exist in conducting such remote user studies. Further, it can uncover specific strengths and weaknesses within the platform itself.

In this two-week experiment, the *online dashboard* with exergame tracking is introduced as the intervention throughout the study. For *focus B: Gamification techniques for exercise motivation*, the experiment could potentially display if there are any motivational effects related to the tracking provided by the *online dashboard*.

Moreover, the goal was to always try to capture participants thoughts on the platform they were testing. Their subjective opinions in combination with the data from game-logs could potentially discover whether there were any effects in terms of engagement caused by the introduction of the *online dashboard* as the intervention.

4.2.1 Study Design

This practical user study aimed at investigating whether the *online dashboard* would have any effect on the engagement of participants, as stated above. However, there are different ways this intervention can be introduced into the study, and I did consider a few options when designing the study experiment.

Original Plan

The plan was originally to split participants into a *Control Group* and a *Treatment Group*. These groups would be split like this:

[3] facebook.com
[4] reddit.com/r/ntnu/comments/mp9vfu/research_survey_experience_with_workout_apps/
[5] reddit.com/r/SampleSize/comments/mp9kyd/academic_exercise_and_workout_tracking_apps/
[6] innsida.ntnu.no
[7] discord.com
[8] ibm.com/analytics/spss-statistics-software

Control Group

Doing tasks in the VR application.

Treatment Group

Doing tasks in the VR application. Introduced to the online dashboard in the middle of the study.

$$
\begin{aligned}
CG &: O_1 \xrightarrow{\hspace{8cm}} O_2 \\
TG &: O_1 \xrightarrow{\hspace{2.5cm}} X_{\text{onlineDashboard}} \xrightarrow{\hspace{2.5cm}} O_2
\end{aligned}
$$

Figure 38: Original Plan: Pretest - posttest experiment design

Figure 38 also displays this original study design. However, when the study experiments were conducted, there were not enough participants to justify splitting them into two separate groups. A few adjustments had to be made to the study design.

Final Study Design

Instead of splitting groups into two, I decided to make all participants partake as part of the *Treatment Group*. As the treatment group already was designed to have the treatment introduced halfway through, this would mean that participants themselves could be their own control group. The idea is to see if their behaviour or engagement changes after the intervention is introduced, which would still be possible to see.

It is relevant to acknowledge that this brings up some other biases which are harder to discover. One example is that participants may feel more engaged at the beginning of the user study, as the platform they are testing out is new to them. This could make it somewhat harder to see whether the introduction of the online dashboard has any motivational effects, but still be possible.

The final study design had this participant group:

Treatment Group

All participants will do tasks in the VR application. Everyone is introduced to the online dashboard midway through the study.

$$
TG : O_1 \xrightarrow{\hspace{2.5cm}} X_{\text{onlineDashboard}} \xrightarrow{\hspace{2.5cm}} O_2
$$

Figure 39: Final Study Design: Pretest - posttest experiment design

4.2.2 Experiment Schedule

From the final study design, I created a more detailed schedule with everything that was planned for the two-week study. This displayed when participants would be introduced to tasks and a high-level overview of what is included.

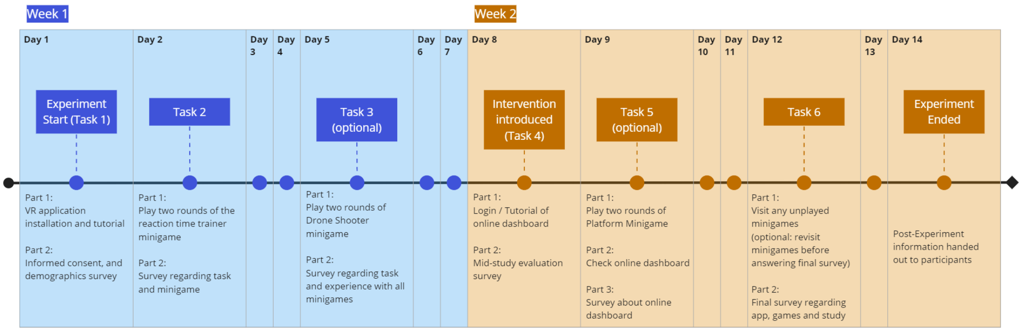


Figure 40: Experiment Schedule for the Two-Week User Study

As fig. 40 displays, there was a total of *6* tasks for participants to do over the two weeks. I estimated that each task would take around ∼15 minutes to complete. The estimate was based on a 1-day pilot stress test (see section 4.2.4 below) that was conducted prior to the experiments.

As the tasks were quite time-demanding for participants, I decided to go divide up tasks such that there were *2* optional tasks, and *4* obligatory tasks. This served two purposes: Firstly, it could make it easier for people to say yes to partake in a study when it was not as demanding. Secondly, it could be interesting to see whether the introduction of the intervention could affect people's engagement in the study to partake in optional tasks.

4.2.3 Task Sheets

I created task sheets for each task of the study. These included explanations on the task, with screenshots and images to further elaborate on details to participants. The task sheets that were distributed throughout the study can be seen in this repository: github.com/akerholten/MSc-Thesis/tree/main/Two-WeekExperiment/TaskSheets

It was important to make sure that these task sheets would provide sufficient information to participants as there would be less room for direct facilitation from me, as it was a remote user study. I previously discovered in the literature review that the lack of facilitation can cause issues to people's efficiency with tasks (see section 2.3.2). To mitigate this, I paid effort towards making sure that the tasks were well described and with visual explanations.

Additionally, I also told every participant that I was available if they would wish to have a 1-on-1 session with me, both before and during the study. This could give participants a better feeling of being *taken care of* by me as a researcher, in case they would feel confused or have any concerns. Yet, I did not want to push this on participants. I did also discover from the literature study (also in section 2.3.2) that participants could feel less stress and more relaxed from participating in a study without having someone observing them. Thus, leading to the decision of letting the participants decide whether they wanted this 1-on-1 facilitation or not. In the end, there were no participants that wanted a direct 1-on-1 interaction. If there were any concerns, they took this up with me through text communication.

Questionnaires

For every task, there was always a questionnaire for participants to fill out afterwards. This would collect data related to the task or overall *ExerIsland* platform. It also was a way to make sure that participants could input some qualitative data and thoughts. Further, if the participants had any concerns, they could bring them up while remaining anonymous. Questionnaires used throughout the study can be reached here:

github.com/akerholten/MSc-Thesis/tree/main/Two-WeekExperiment/Questionnaires

4.2.4 1-Day Pilot Stress Test

Prior to starting the full two-week experiment period, I wanted to make sure that the tasks seemed manageable for participants to complete. To ensure this, I conducted a "*stress test*" pilot study with one person. This was a 1-on-1 session with me, where I could collect qualitative data by observing how the person performed tasks and allowing the person to give direct feedback to me throughout this test.

This pilot stress test was a way to do some quality assurance before starting the whole study. As it would be much harder to implement any changes midway through the study. This test uncovered mostly simple flaws in task explanations that had to be changed. In addition, it made me make some minor changes to certain tasks that seemed to take a too long time to complete.

4.2.5 Querying for Participants

This study required subjects to dedicate quite some time for participation, and it was a requirement that participants had personal access to VR equipment on PC. Because of these restrictions, I had to spend quite some time querying for participants for this experiment.

When I queried for participants in this study, I did it similarly as for the *questionnaire study*. I reached out to people in my own network and on channels related to NTNU Gjøvik. Additionally, my *questionnaire study* contained an information note about this two-week VR experiment. This meant that people could contact me if the extensive study experiment seemed interesting to them.

I distributed invitations to the two-week experiment on these platforms:

- Facebook [9]
- NTNU Gjøvik's Discord [10] channels

There was a total of *7* people (aged 18-30 years old) who participated and completed the whole user study.

4.2.6 Data Analysis

I used Jupyter Notebook [11] with python to extract certain data points from the data collected through the ExerIsland application. Due to the structure of the game logs having several layers with *Sessions->Activities->Metrics* (see data structure in section 3.2), the data had to be worked a bit with to get a two-dimensional data structure for outputting to a Comma-separated values (CSV)-file for further data analysis.

[9] facebook.com
[10] discord.com
[11] jupyter.org

Then, I synthesized the data from game logs together with data from the questionnaires conducted in the study. In the end, I had a rather substantial CSV file containing all the data collected throughout the study. I used SPSS [12] for the final analysis of the combined data and graph-plotting. Results from data analysis can be seen in section 5.2.

[12] ibm.com/analytics/spss-statistics-software

5 Experiment Results

5.1 Questionnaire Study on Exercise Trackers

In the initial phases of the analysis of this online questionnaire study, I did a quite large *independent samples t-test* [67] to test whether there were any statistically significant findings based on the data from the *94* participants. Results of this online questionnaire study are mainly relevant to *focus B: Gamification techniques for exercise motivation*.

The independent variable I was operating with was *exercise tracker usage*. I collected whether the participant was actively using, previously had used, or never had used any exercise trackers before. This information could be extracted from several sources of data in the questionnaire. The one used for most parts in the data analysis is the question displayed in fig. 41:

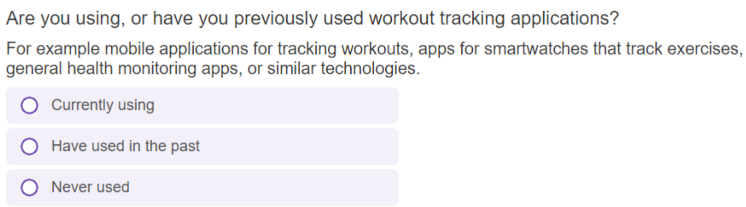


Figure 41: Collecting Independent Variable - Are you using exercise trackers?

Furthermore, I put very many of the dependent variables collected through the questionnaire into the initial *independent samples t-test*. This gave quite good initial indications of the captivating findings of the data or where any statistical significance could be proven or not.[1]

5.1.1 Weekly Exercise Hours

The first thing I saw when comparing the groups of the independent variable was the difference in the reported amount of workout hours per week.

Group Statistics

	Using workout apps or not	N	Mean	Std. Deviation	Std. Error Mean
Weekly workout hours	Currently Using	50	4,96	3,741	,529
	Previously used or never used	44	3,07	2,627	,396

Figure 42: Group Statistics: Workout hours comparison between Currently using trackers and Previously used or never used

[1]Images displaying the initial T-test can be seen in this repository: github.com/akerholten/MSc-Thesis/tree/main/OnlineQuestionnaire_ExerciseTrackers/Analysis

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Weekly workout hours	Equal variances assumed	3,052	,084	2,800	92	,006	1,892	,676	,550	3,234
	Equal variances not assumed			2,862	87,861	,005	1,892	,661	,578	3,205

Figure 43: Independent Samples T-Test: Workout hours comparison between Currently using trackers and Previously used or never used

The group statistics and T-test in figs. 42 and 43 displays that the participants that currently is using exercise trackers seem to report a higher average mean in terms of *workout hours* per week.

However, the p-value (*Sig.* value) in Levene's Test is .084, which means that the variance between the groups is not statistically significant. Albeit, the result of the t-test (sig. 2-tailed) points towards a low p-value of .006. This indicates a statistically significant difference in the mean values between the two groups. To further check this data, I went a bit deeper and tried testing with *Daily exercise tracker usage* as an independent group variable to see whether there were any even more statistically significant findings with that.

Group Statistics

	When using workout tracking apps, how often did you use it?	N	Mean	Std. Deviation	Std. Error Mean
Weekly workout hours	Used exercise tracker daily	37	5,46	4,100	,674
	Not using daily or never used	57	3,18	2,479	,328

Figure 44: Group Statistics: Workout hours comparison between Daily using trackers and Not using daily or never used

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Weekly workout hours	Equal variances assumed	6,337	,014	3,368	92	,001	2,284	,678	,937	3,631
	Equal variances not assumed			3,046	53,192	,004	2,284	,750	,780	3,788

Figure 45: Independent Samples T-Test: Workout hours comparison between Daily using trackers and Not using daily or never used

Figures 44 and 45 displays a statistically significant ($p < .004$) mean average of the weekly workout hours. I must mention here that this group represents a very active user group of exercise trackers, and it is expected to see a quite notable difference in workout hours. The group that used exercise trackers daily did report a mean difference of more than 2 hours more of physical exercise each week. I should also add that the 95% CI indicates that there is a distinction, but somewhere between [.780, 3.788] hours. This means that the data points in that direction but can not pinpoint exactly how many hours difference there is between the groups.

Even though users actively using workout trackers report a higher average amount of workout hours each week, it is hard to know whether this is a causal connection or just a correlation. It is quite likely that people that generally exercise more, are inherently more likely to use exercise trackers on a daily basis.

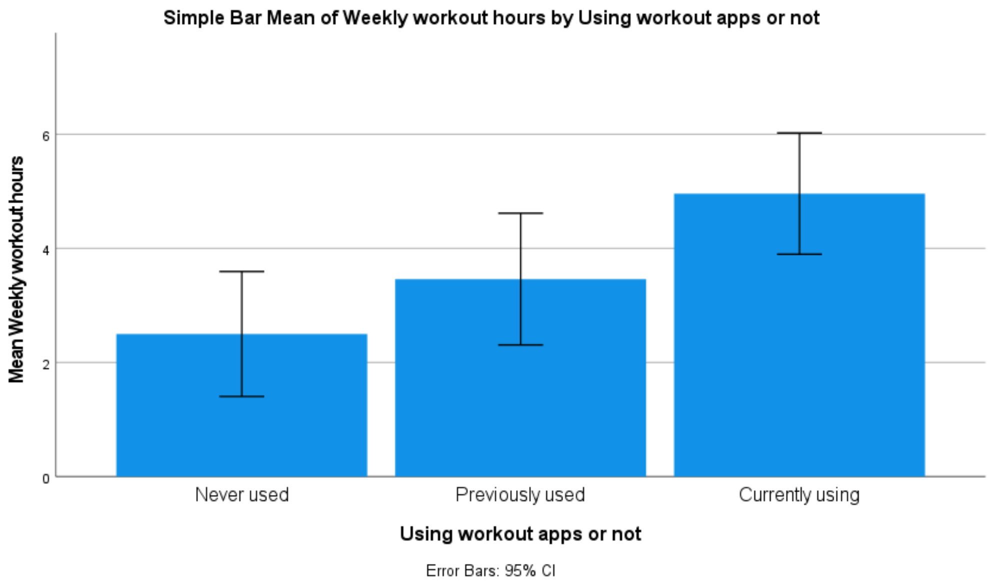


Figure 46: Graph: Weekly workout hours compared with exercise tracker usage

Figure 46 is a graphical display of the current findings I have described here. In the graphs displaying results, the *independent variable* is split into *three* categorizations:

1. **Never used** exercise trackers
2. **Previously used** exercise trackers
3. **Currently using** exercise trackers

Interestingly, the different grades of having used exercise trackers correlate rather well with the weekly workout hours. I must also mention that the 95% CI bars display the data could potentially change quite a lot with a larger sample size, but they display a good indication.

To further elaborate on the influence on workout hours caused by exercise trackers, we can take a look at participants self-reported data relevant to this:

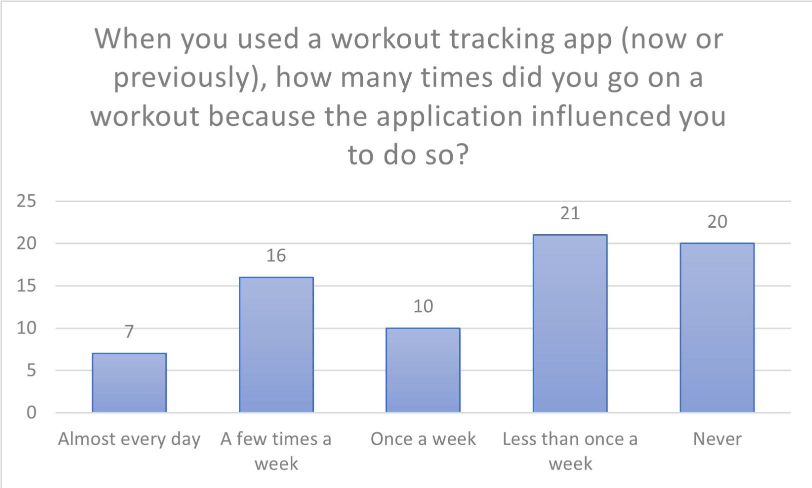


Figure 47: Graph: Self-reported responses on how often workout tracker influenced subject to exercise

The graph displays quite varied data concerning whether they themselves have observed that the exercise tracker has influenced them to do physical activity. However, out of the *74* responses to this question, only *20* responded that their tracker had an influence on them exercising. I find these results to be quite positively oriented. At the same time, it is still based on personal reflection on own workout habits, so it can be hard to verify the accuracy of this reported data.

5.1.2 Motivations to Exercise

We have now looked at how exercise tracker usage correlates with reported workout hours. To get a further understanding of this correlation, I wanted to investigate people's reported motivations to do physical activity.

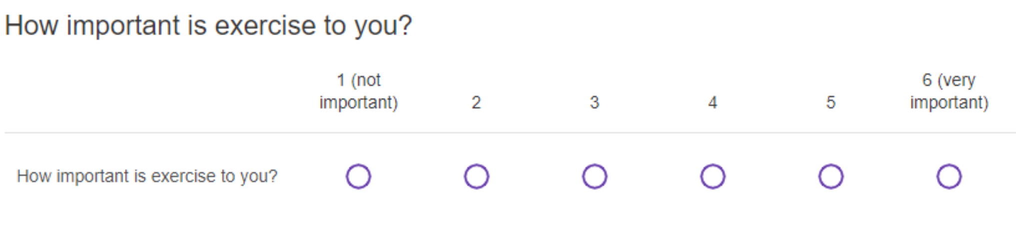


Figure 48: Question from Questionnaire - How important is exercise to you?

Participants rated how important exercise was to them personally, and the results from this question interestingly display a positive oriented-result from all groups:

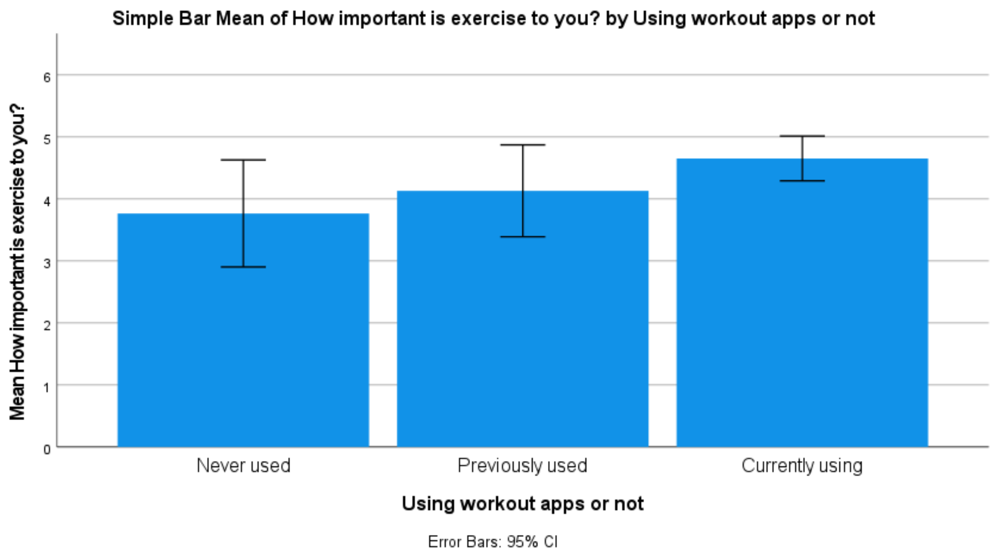


Figure 49: Graph: How important is exercise to you?

The results displayed by this graph (fig. 49) is actually quite interesting. Despite the differences in workout hours displayed above (section 5.1.1), the findings show that the average level of importance of exercise is pretty positively oriented. The data displays a much smaller difference compared to the average workout hours per week (see fig. 46). An assumption to make from this data is that people seem to evaluate exercise to a certain level of importance, but it does not map directly to the amount of exercise people do.

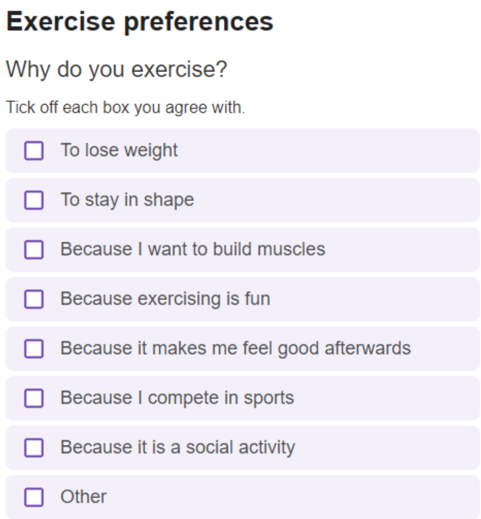


Figure 50: Question from Questionnaire: Why do you exercise?

Based on the initial T-Test, the data seemed to have quite some differences in the groups using and not using exercise trackers from the question above (fig. 50). These data-points are either set to *true (1)* or *false (0)*.

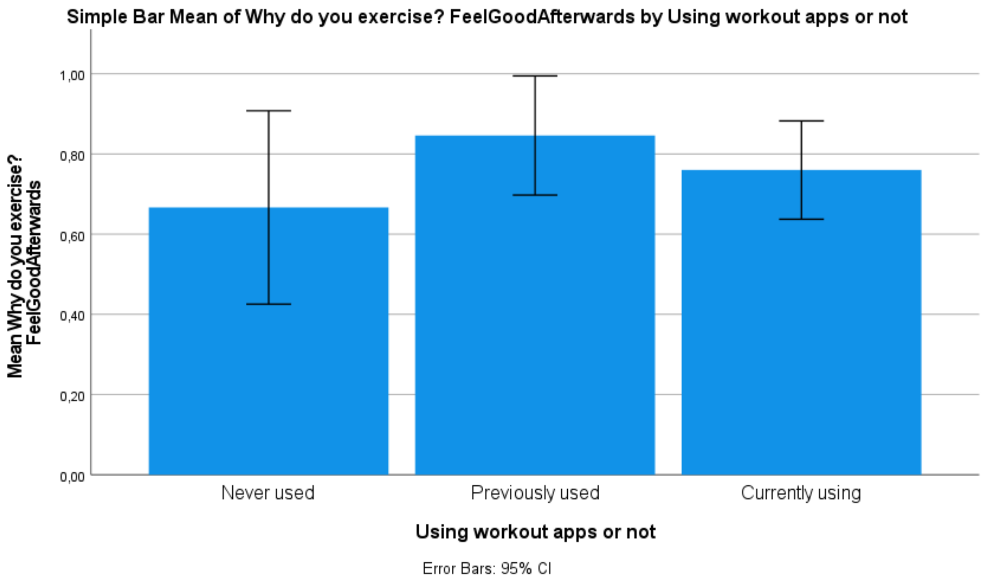


Figure 51: Graph: Why do you exercise? "Because it makes me feel good afterwards"

Concerning people motivation because workout gives them a sensational feeling afterwards, this was one of the more equally reported motives found in the data. As seen in the fig. 51 above, this motivation was reported fairly high across all three categories of exercise tracker usage.

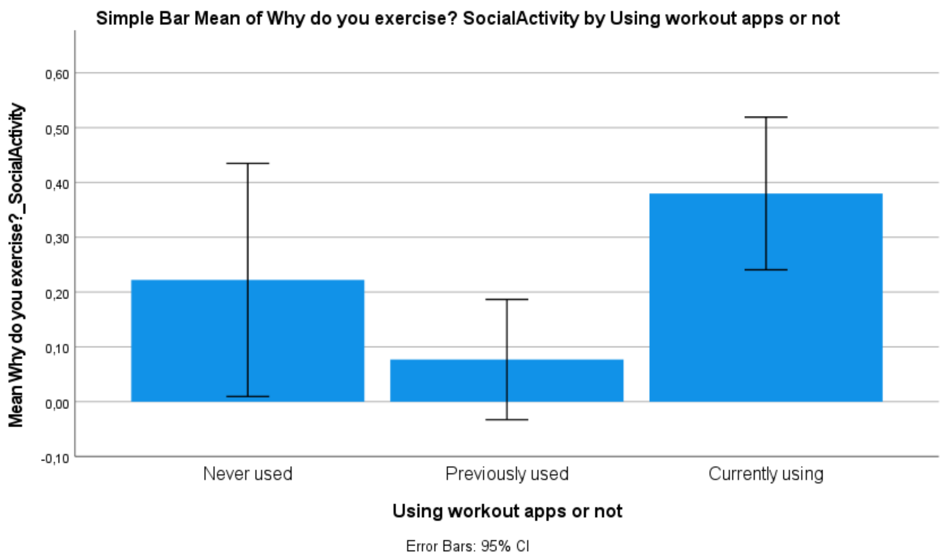


Figure 52: Graph: Why do you exercise? "Because it is a social activity"

Figure 52 displays the average reports of being motivated to exercise because it is a social activity. This is interesting to point out as there is quite a broad variance between all three categorizations. The *never used* category displayed a higher response compared to the *previously used*. However, as the 95% error bars indicate, there is not a large enough sample size to accurately pinpoint this data.

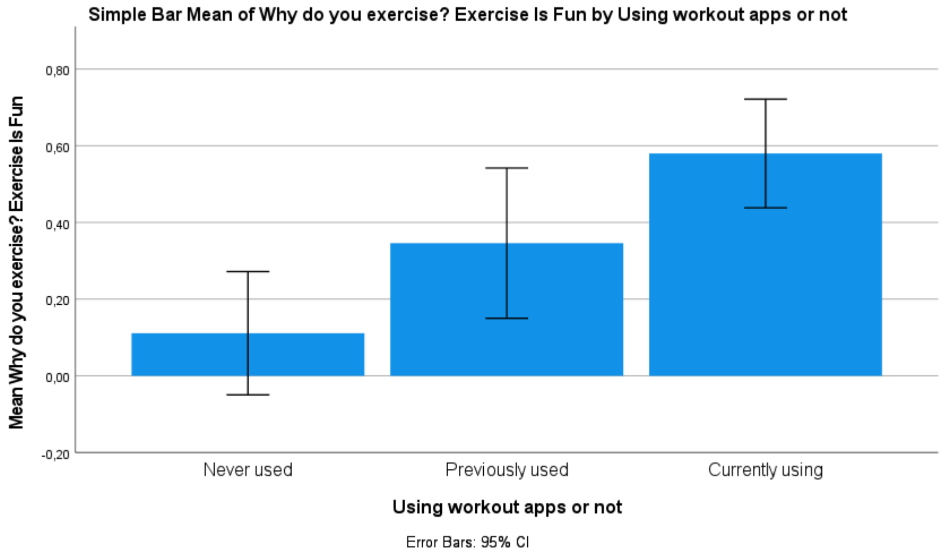


Figure 53: Graph: Why do you exercise? "Because exercising is fun"

Now, one of the most exciting findings, in my opinion, is this self-reported data on whether people exercise because they find the exercise itself to be a fun activity (fig. 53). The error bars in this graph does indicate that some potential variance can be found with a larger dataset. Nevertheless, the results here suggest that there are much more people currently using exercise trackers that finds workout as *fun*.

It is not necessarily possible from this data to confirm if this describes a causal relationship. Though, it does describe a quite significant correlational difference. It might be that the people that have more fun with physical activity also do more exercise. In turn, this might also result in more of these people using exercise trackers. The data is still an indication that the gamification provided by an exercise tracker might affect peoples perceived *fun* of the activity itself. A T-test was also performed regarding this metric from the questionnaire. It displays a statistical significant difference ($p < .001$) between the two compared groups (see figs. 54 and 55 below).

Group Statistics

	Using workout apps or not	N	Mean	Std. Deviation	Std. Error Mean
Why do you exercise? _ExerciseIsFun	Currently using	50	,5800	,49857	,07051
	Previously used or never used	44	,2500	,43802	,06603

Figure 54: Group Statistics: "Exercise is fun" comparison between Currently using trackers and Previously used or never used

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Why do you exercise? _ExerciseIsFun	Equal variances assumed	11,416	,001	3,388	92	,001	,33000	,09741	,13654	,52346
	Equal variances not assumed			3,416	92,000	<,001	,33000	,09660	,13814	,52186

Figure 55: Independent Samples T-Test: "Exercise is fun" comparison between Currently using trackers and Previously used or never used

Participants also rated whether they thought workout tracking software could motivate them to exercise or not. This was a rather direct question and is based on subjective opinion, but it does give a further indication of people's own trust in the technology.

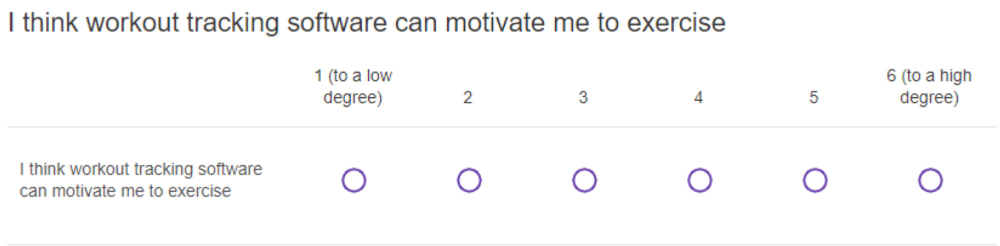


Figure 56: Scale from Questionnaire: "I think workout tracking software can motivate me to exercise"

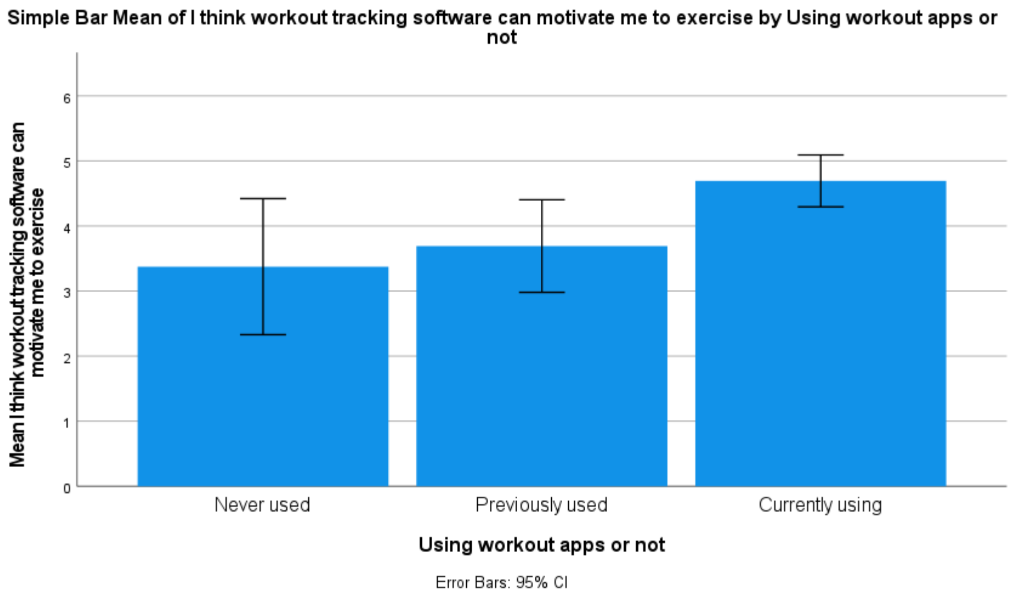


Figure 57: Graph: I think workout tracking software can motivate me to exercise

The results are generally positively oriented. However, there is a trend displaying that the people currently using exercise trackers report a higher belief in its motivational effects.

5.1.3 Importance of Features

Another valuable aspect to investigate from the data was the perceived importance of specific features in trackers. To describe these findings, I will still make use of the *three* categories of tracker-usage, as they do indicate some discrepancies between the groups. These differences are critical to acknowledge as they can shed light on what parts of exercise trackers pulls existing users to these apps. Moreover, clarify what is potentially missing for the audience not using them.

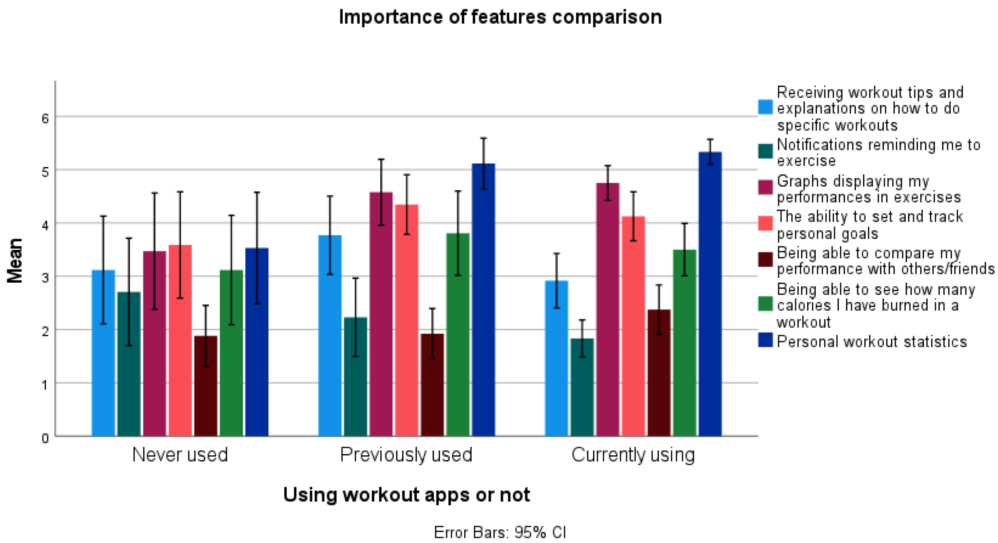


Figure 58: Graph: Importance of features in exercise trackers

We can see that there are some common trends among all three groups. The most prominent feature from the data seems to be *personal workout statistics*, and the score for this feature is a lot higher in groups using exercise trackers compared to the group not using.

Features in regards to *comparing performance with others* seem to be the lowest-rated across the whole board. This feature was only being slightly higher rated in the group that is *currently using* trackers, but still very low.

Some other fascinating findings from this graph is that the groups that currently are not using exercise trackers value *notifications reminding them to exercise* and *workout tips* higher than the group which is using these trackers. A thing to note is that it might be that these subjects are simply exercising less and are self-aware about that. In turn, this could lead to higher scores in these two features. However, it could also mean that the people in this group have not found exercise trackers who fulfill their needs concerning these features. Thus, making them not use any exercise trackers either.

We can see that in the *currently using*-group, the feature of getting notifications to remind them to exercise is the lowest scored. This complements the earlier findings regarding average workout hours each week. It is quite likely to assume that many of the participants that use trackers have good exercise schedules and thus would not see the benefit of having notifications reminding them to exercise.

Overall, the graph in fig. 58 displays an interesting trend towards features deemed more relevant in both current- and potential new users of exercise tracking applications.

5.1.4 What made people quit using the trackers?

Let us move over to another alluring area of findings. There is a vast amount of participants who have reported using exercise trackers but ended up quitting using them. It can be critical to try to understand why they stopped using their trackers if it was due to reasons implied from the tracker itself or something else.

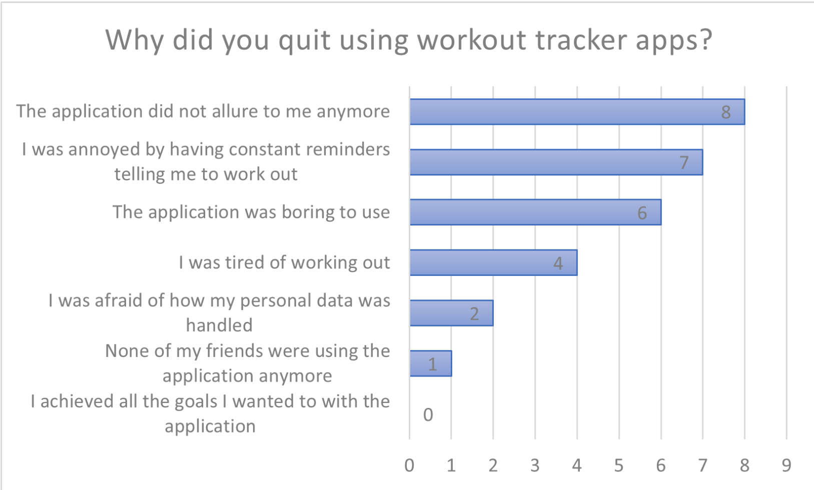


Figure 59: Graph: Why did you quit using workout tracker apps?

Figure 59 indicates that the most common reason people had for quitting using their exercise tracking application was that they got tired of the app itself. Out of *24* respondents, 8 reported that *the application did not allure to them anymore*. This is tightly followed up with 7 people saying that they were *annoyed by having constant reminders from the app*. In terms of data privacy in software, there were only 2 people who reported that they quit using their tracker due to being *afraid of how their personal data was handled*.

In addition to the options above, there were some free-text answers in response to this question. Meaning that people had *other* reasons to quit using their exercise tracker than what was listed above. Some of these can be seen below:

> *I have not quit. I am not actively using. In winter I do not go for hikes. So I will start using it back soon when it is good time to go hiking.*
>
> *- Respondent 1*

> *My workout routine changed. I used tracking apps for long cycling trips. Now I run without a cellphone. Might use exercise apps in the future, during vacation and longer trips.*
>
> *- Respondent 2*

> *My app was specifically for the gym, and the gym is closed due to covid.*
>
> *- Respondent 3*

I lost my watch.

- Respondent 4

I put my phone back in my native language and now the app does not work as good.

- Respondent 5

One can see that these free-text responses indicate that some people that stopped using their tracker did this due to external reasons and not because of the tracker itself. One indication to take from this though is accessibility in software. Respondent 5 reported that putting his phone back in his native language made the application not work as good. Issues like these should be managed from a software engineering perspective. The application should not render itself unusable due to a simple change like native language on the phone.

Let us take a look at peoples subjective opinions on whether they exercise more or less after they quit using their tracker:

Figure 60: Graph: Since you stopped using your workout tracker, how much do you now exercise?

Figure 60 displays a quite mixed result. A total of 9/24 respondents reported now exercising more than previously, while 15/24 reported less than when they used their tracker. The findings here do complement the findings of average workout hours quite well (see section 5.1.1). The results are oriented towards a lower amount of exercise after stopping using their tracker. It is worth noting that most responses are close to the *middle* area, indicating a minimal change.

5.1.5 Weekly Videogame Hours

Another striking finding discovered during data analysis was the indicative difference in the number of hours people spend on videogames across the categories defined. Although this is somewhat outside of the focus area, I wanted to just quickly describe this finding as it was such an obvious difference over the data.

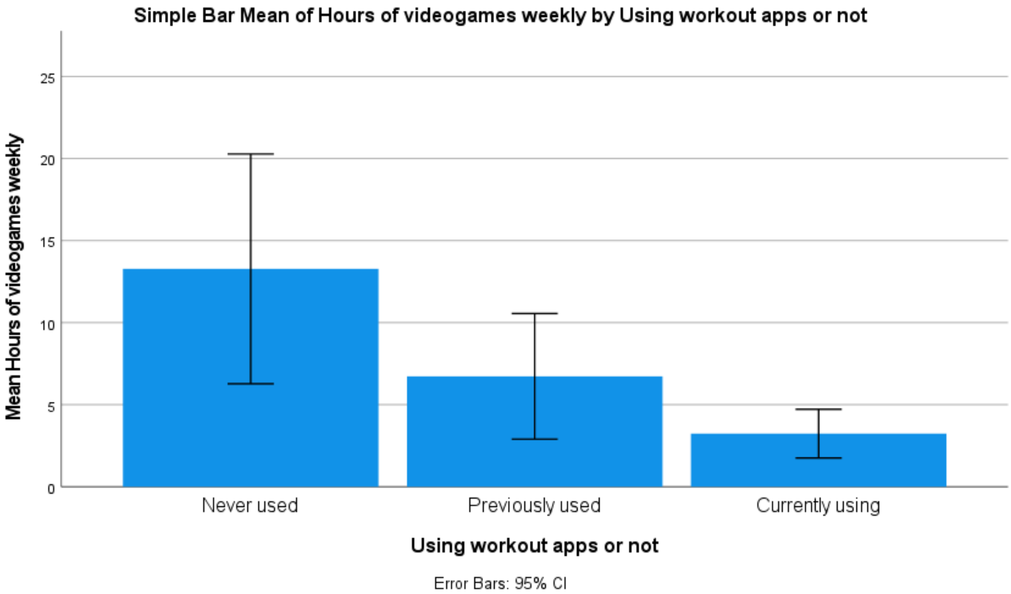


Figure 61: Graph: Hours of playing video games each week

I wanted to bring up this result as a goal of this thesis is to investigate whether gamification can increase peoples motivation for physical activity. In this regard, this specific graph displays that people that do not use exercise trackers spend more time each week playing videogames on average. It is uncertain whether this actually means that these people are less prone to liking these gamifications in the exercise trackers or if it is just the fact that people that play more videogames tend to exercise less. The findings do at least display a fairly obvious correlation in the data.

5.2 Two-Week Experiment with Exergame Platform

For this experiment, I wanted to investigate the engagement of users in the application before and after introducing the intervention with the online dashboard. In addition, the findings here describe how participants in the user study evaluated the platform. This evaluation is for how functional implemented features were, but also how valuable they found different features of the platform.

5.2.1 User Engagement - Week One & Week Two

Based on the logs collected from VR gameplay, I could synthesize some alluring parts of the data together. The idea was that this data potentially could objectively describe whether users were more or less engaged in the second week of the study.

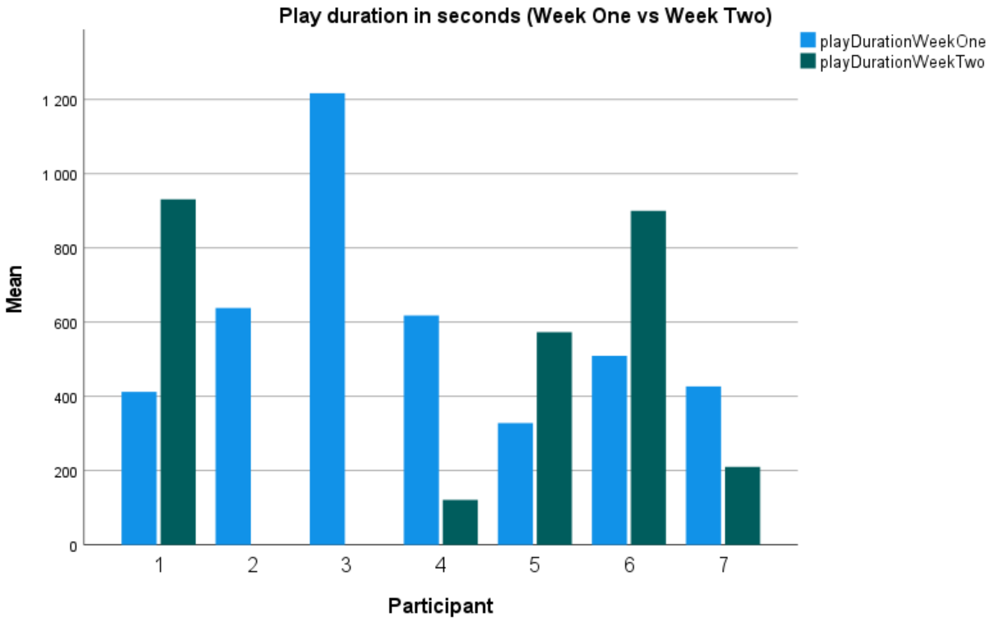


Figure 62: Graph: Play Duration - Week One & Week Two

Figure 62 displays the total play duration (in seconds) of *week one* and *week two* data for all participants. At first sight, it seems to give a perspective that the tendency is towards a much higher play duration in week one. There is indeed on average a higher play duration in week one. Albeit, there are a few important things to realize with this fact:

First of all, there are 2 participants (participant 2 and 3) that seemingly only played minigames in week one of the study. Secondly, a trend for these 2 participants is that they have a higher average play duration in week one than all other participants. This will further skew the data towards more engagement in week one than week two. Due to the nature of a remote user study, it is hard to tell whether this data is correct or if the users had any technical issues on their end that hindered logs from being sent to the backend. It might also be that the data is totally correct, and these participants just decided to do much more in the first week and then not in the second week.

Let us look at the data from the 5 other participants (1, 4, 5, 6 and 7) that used the VR application in both weeks of the study. By investigating these subjects the trend seems considerably different in fact. From these 5 people, we see that *3* (1, 5 and 6) had quite an increased play duration after the intervention was introduced. The remaining 2 participants (4 and 7) display a decreased play duration in week two.

Overall, the data points to mixed results. Based on the data collected through logs it seems like 2 participants never played any VR games in the second week, 3 participants had an increased play duration, and the remaining 2 displayed a decreased play duration in the second week.

Group Statistics

	WhichWeek	N	Mean	Std. Deviation	Std. Error Mean
totalSessionCount	Week Two	7	1,29	,951	,360
	Week One	7	1,86	,690	,261
activityCount	Week Two	7	4,00	4,359	1,648
	Week One	7	5,86	2,795	1,056
totalPlayDuration	Week Two	7	390,71	407,273	153,935
	Week One	7	592,71	297,104	112,295
totalArmMovement	Week Two	7	680,86	793,458	299,899
	Week One	7	941,29	546,995	206,745
totalAvgScore	Week Two	7	2337,85	2515,960	950,944
	Week One	7	3423,10	1669,454	630,994
RTT_Activity_Total	Week Two	7	1,29	2,215	,837
	Week One	7	3,43	1,397	,528
DroneShooter_Activity_Total	Week Two	7	1,14	1,676	,634
	Week One	7	1,86	1,345	,508
Platform_Activity_Total	Week Two	7	1,57	1,272	,481
	Week One	7	,57	,976	,369

Figure 63: Overall Group Statistics - Comparing users with themselves (Week One vs. Week Two)

To further inspect the data, I created a dataset that I could use to compare users data with themselves. I did this by splitting their logs from week one and week two into two separate objects in the dataset. The idea was that I could use this to emulate a *Control Group* versus *Treatment Group* comparison. Figure 63 displays the overall statistics. The trend points toward that participants were more active in week one of the study. However, as previously pointed out, two participants had no logs of doing activities in week two of the study.

If we filter out these 2 participants, the data points a bit differently:

Group Statistics

	WhichWeek	N	Mean	Std. Deviation	Std. Error Mean
totalSessionCount	Week Two	5	1,80	,447	,200
	Week One	5	2,00	,707	,316
activityCount	Week Two	5	5,60	4,159	1,860
	Week One	5	4,80	1,924	,860
totalPlayDuration	Week Two	5	547,00	376,758	168,491
	Week One	5	458,80	109,744	49,079
totalArmMovement	Week Two	5	953,20	787,316	352,098
	Week One	5	711,60	433,998	194,090
totalAvgScore	Week Two	5	3272,99	2381,011	1064,821
	Week One	5	3764,14	1891,394	845,857
RTT_Activity_Total	Week Two	5	1,80	2,490	1,114
	Week One	5	3,00	1,000	,447
DroneShooter_Activity_Total	Week Two	5	1,60	1,817	,812
	Week One	5	1,80	1,483	,663
Platform_Activity_Total	Week Two	5	2,20	,837	,374
	Week One	5	,00	,000	,000

Figure 64: Overall Group Statistics - Comparing users with themselves (Filtered out participants inactive in week two)

The measurements of *activityCount*, *totalPlayDuration* and *totalArmMovement* all point towards a higher average in the second week when we have filtered out the two participants who did not use the VR application in the second week.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
totalPlayDuration	Equal variances assumed	8,517	,019	,503	8	,629	88,200	175,494	-316,489	492,889
	Equal variances not assumed			,503	4,674	,638	88,200	175,494	-372,554	548,954

Figure 65: T-Test - Play duration week one compared with week two (Filtered out participants inactive in week two)

Figure 65 displays the results of a T-Test in regards to the play duration in week one versus week two. *Levene's test for equality of variances'* low sig. value indicates that the variance in play duration is statistically different between the two groups, which means we should look at the *'equal variances not assumed'* for the t-test. However, by looking at the t-test for equality of means, we see that the p-value (Sig. 2-tailed) is way too high (.638). This means that we can not reject the null hypothesis. In other words, there was not sufficient evidence to prove that the intervention (online dashboard) had an effect on the participant group.

5.2.2 Users' Evaluation of the VR Application

Features of the VR Hub

Participants rated different features within the VR application on a scale from 1 (very bad) to 6 (very good). Figure 66 below displays the combined results from all participants for these features.

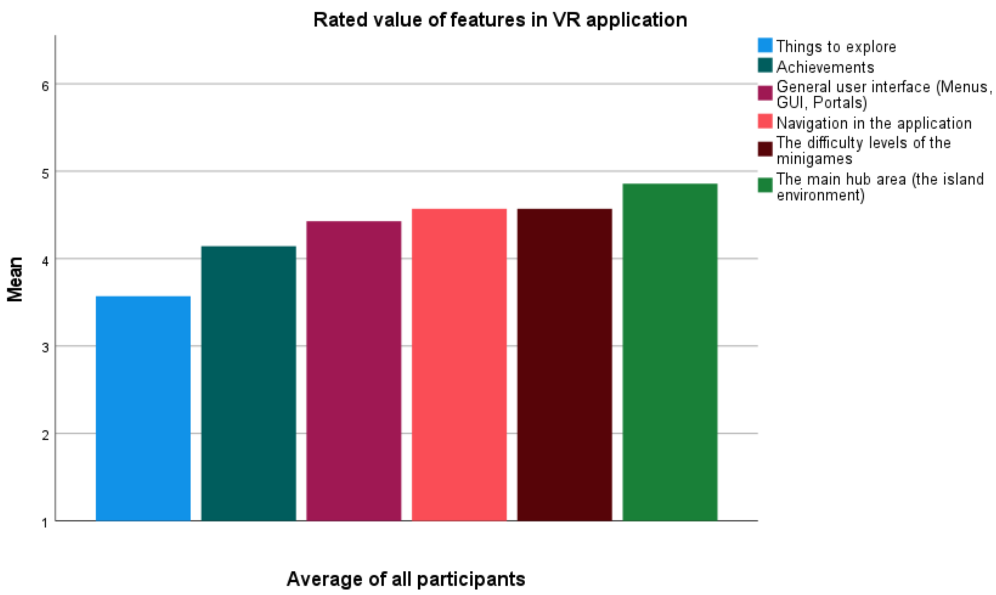


Figure 66: Graph: Rated features of the VR application

The feature that received the lowest reported score was *'Things to explore'*. I did previously mention in section 3.4.4 that the application was somewhat limited in its features for the *explorer* player type in Bartle's taxonomy. [63] With this graph, it is displayed that it corresponds rather well with my personal reflections on the platform as well.

Apart from that, the overall features of the VR application were quite positively oriented. Though, there is room for improvement in most categories. The feature participants rated the highest within the VR application was the main hub environment. In regards to the hub, there were also a few text comments from participants further emphasizing this:

> *The games have been fun and engaging, and the hub concept seems to work quite well.*
>
> *- Participant 4*

> *Great work overall. The graphics are nice and calm. The minigames are good. Hopefully, some of the data & feedback will be useful and help your development further. I hope to see more of the game someday.*
>
> *- Participant 1*

> *The hub is good use of VR!*
>
> *- Participant 5*

> *The game world is pretty, and sunny, reminds me of summer and happy times. Nice and funny minigames.*
>
> *- Participant 7*

Potential Improvements for the VR Hub

I asked participants to provide feedback concerning what they were missing or would like to see improved within the application. Below are some of the comments from participants in regards to this:

> *Exploring and interacting with world objects. Level progression in the minigames. Leaderboard - being able to compare yourself or versus others in-game.*
>
> *- Participant 1*

> *A challenge and depth to the games.*
>
> *- Participant 2*

> *I would like more interaction in the hub, things to play with etc. I would also like to see a more general theme throughout both the hub and the minigames.*
>
> *- Participant 4*

> *Would love to see the hub fleshed out a bit and maybe integrated with the web application somehow. Would also like the menus to maybe be more interactive by making the buttons physically interactable and 3D.*
>
> *- Participant 5*

> *Higher fidelity minigames, as in tighter mechanics. Depth in the minigames, as in more mechanics to consider at once. More content, maybe random generation of levels could help with this, like a floor plan with rooms like a roguelike where exercise activities were integrated in the combat/puzzle-solving (maybe to open doors).*
>
> *- Participant 6*

> *It would have been fun if you could play against others inside the platform and see the results of the others (my friends)'s result versus mine.*
>
> *- Participant 7*

The comments point out several points of improvement for the VR application. Some participants indicated lack of features that are important to a *socializer* or *killer* as a player type. [63] In regards to the missing possibility of comparing their performances with others. Participant 3 and 4 reported that they would like to see more interactivity within the game hub itself, creating more of a play space of things to interact with. Another overall trend from these comments is that the games and hub itself were lacking some depth. Some of the minigames could appear to be too *simple*, which in turn could make them less captivating over time.

Minigames

Participants rated the three different minigames from 1 (very bad) to 6 (very good).

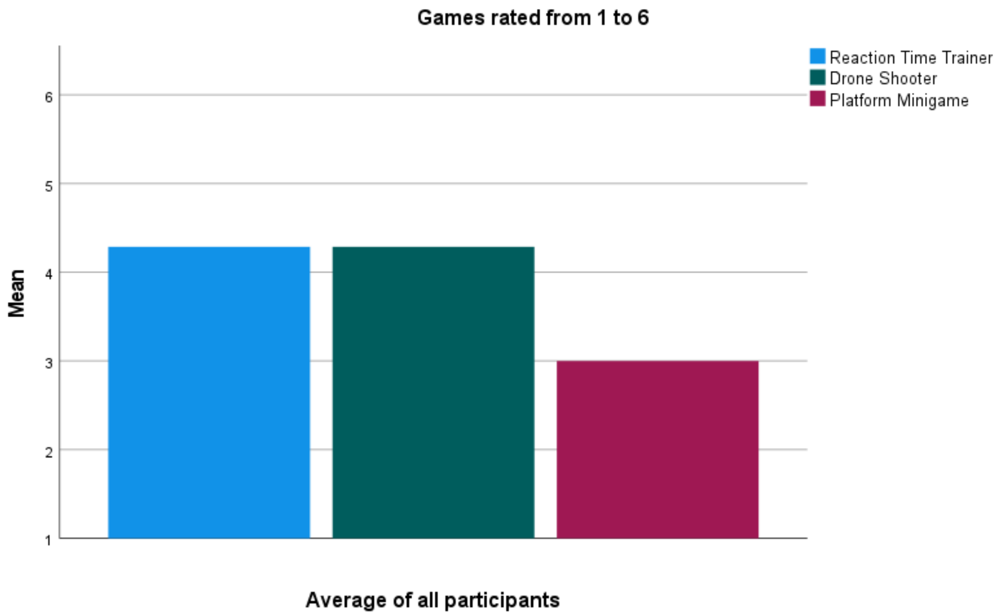


Figure 67: Graph: Rated minigames in VR app

Figure 67 displays the combined result in regards to the rating of the minigames by the participants. *Reaction Time Trainer* and *Drone Shooter* received similar, positively-oriented scores. Platform Minigame, on the other hand, received a slightly worse score, pointing more towards being a slightly *bad* experience.

There were also several comments in regards to the *Reaction Time Trainer* minigame, with several participants indicating this minigame to be very engaging in its gameplay and uniqueness:

> *Extreme was very engaging, and I would like to play it more on this difficulty if I had time!*
>
> *- Participant 5*

> *The game was fun and interactive. The music, haptic feedback, and increasing difficulty made the game quite addictive in a way. I felt inclined to continue playing to increase my score on the harder difficulties.*
>
> *- Participant 4*

> *It was quite fun for how simple it is. I got engaged and wanted to hit them quickly. Though locating which had activated got a little hard as they were out of view, having unique sounds/pitch to each bubble might help. The game was fun, but lacking depth. Too simple. It would not keep me entertained for very long. I enjoyed the experience, but wouldn't be coming back to it.*
>
> *- Participant 6*

There were also comments from people on why they disliked the *Platform Minigame*. Indications are towards that the game feels less polished and that it lacks *design*. Participants indicate that the obstacles spawning seem too arbitrary to engage them in the activity:

> *I feel like the platform mini-game is too "gamey" and the layout of obstacles and target objects feels very arbitrary and in contrast with i.e beat saber does not promote matching the song beat rate.*
>
> *- Participant 5*

> *Platform game has bad collisions or maybe the hands are too small? Does not feel reliable. Additionally, the node spawning in platform does not often place two nodes for one of the hands.*
>
> *- Participant 6*

Short-Term & Long-Term Entertainment

Participants rated how the short-term and long-term entertainment value they felt with the application. Figure 68 below displays the results from each individual participant in this aspect.

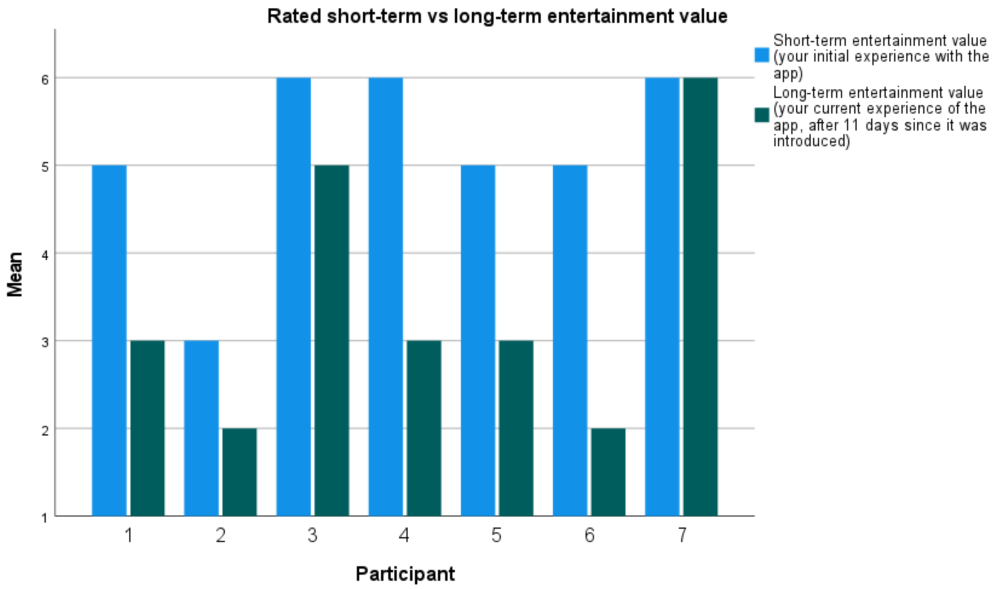


Figure 68: Graph: Rated short-term versus long-term entertainment value of application

The results here indicate a similar trend as seen from comments and data inspection above. Participants generally rated their short-term entertainment value of the application relatively high, with most participants rating it as 5 or higher. Only one participant rated short-term entertainment value lower, at a rating of 3.

However, if we look at the long-term entertainment value, we see that the common trend is not as positively oriented. Overall the long-term entertainment value is rated relatively low for most subjects. Almost all participants (6/7) did also rate a lower long-term entertainment value than short-term entertainment value.

5.2.3 Users' Evaluation of the Online Dashboard

In the second week of the study, I collected subjective data from participants on their thoughts on the online dashboard. Data and findings related to the online dashboard are described below.

Rated Value of Features

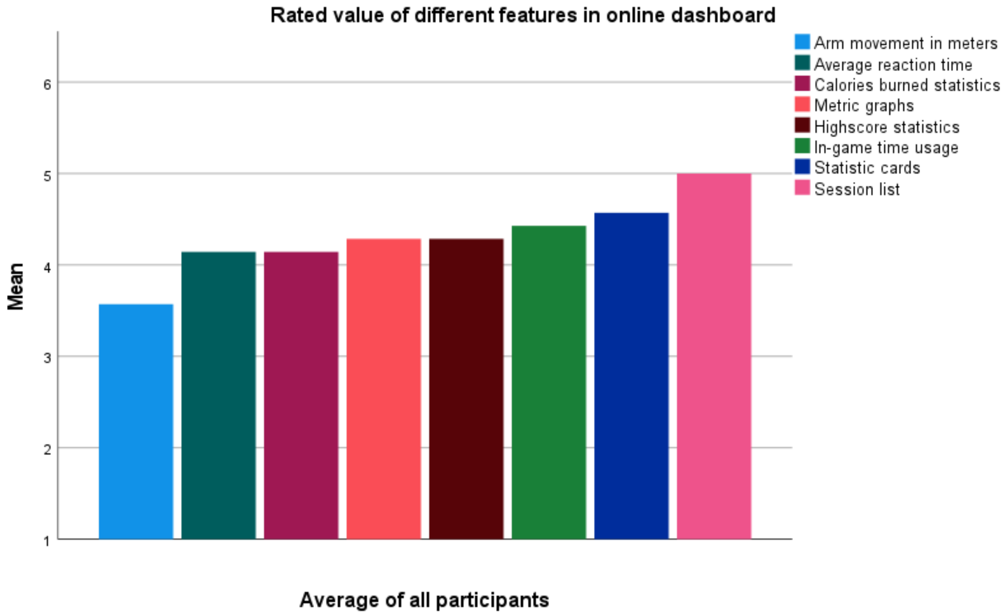


Figure 69: Graph: Rated value of the different features within the online dashboard

Figure 69 displays the participants combined rating on the different features within the online dashboard. The least relevant element of the application seems to be the statistic displaying *total arm movement*. The highest-rated component is the *session list* which gives an overview of all the sessions and activities they have performed.

Generally, the value of features presented within the online dashboard seems to be rated quite positively by the subjects in the study. All features, except *arm movement statistic*, were evaluated above *4* in its score on the 1-6 Likert scale.

Subjective Opinions on Online Dashboard

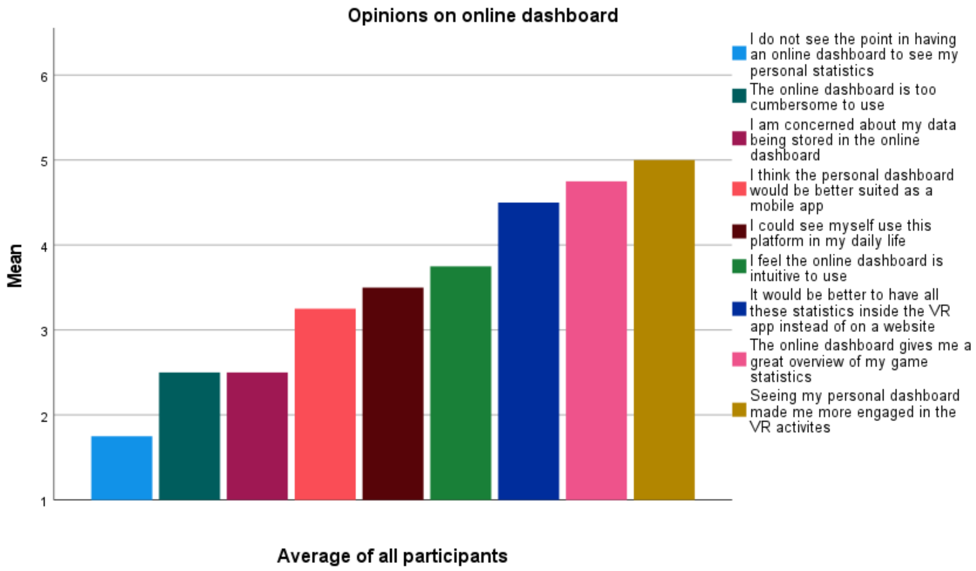


Figure 70: Graph: Opinions on online dashboard rated from 1 (not true) to 6 (very true)

In the fig. 70 above, several statements about the online dashboard have been rated on a 1-6 Likert scale from participants. In my personal opinion, the most interesting points to look at in this graph is the *3* right-most columns. First of all, the average rating on the statement "*Seeing my personal dashboard made me more engaged in the VR activities*" is rated to a *5*. This an interesting finding concerning focus B: *Gamification techniques for exercise motivation*. It indicates that the users personally felt some extra drive towards engaging in the VR activities after using the personal dashboard.

Further, the second-highest rated statement is "*The online dashboard gives me a great overview of my game statistics*". This indicates that users felt that the online dashboard fulfilled its purpose of providing a statistical overview well.

Next, the third-highest rated statement is "*It would be better to have all these statistics inside the VR app instead of on a website*". Users indicated it relevant to have similar statistics available in VR as well, or solely in the VR application.

All of these findings are further highlighted by comments given by participants regarding the online dashboard:

> *The statistics on ExerIsland.com is a great addition to the game. It made me boot up the game one more time! Perhaps more of the statistics should be in-game to make the players push to improve their score, get higher calories burned etc. The score in-game felt a bit boring as your only progression in-game, but if I had seen my calories burned, water bubbles hit etc I would be more engaged to play another round.*
>
> *- Participant 1*

> *The data shown is quite interesting, and the app is well made.*
>
> *- Participant 5*

Can't believe I've moved my arms 2.5km in this test so far. Nice to get an overview of how I perform.

- Participant 6

It is nice to see my results after a VR game. It makes me more motivated, and I want to hit a new record.

- Participant 7

Potential Improvements of the Online Dashboard

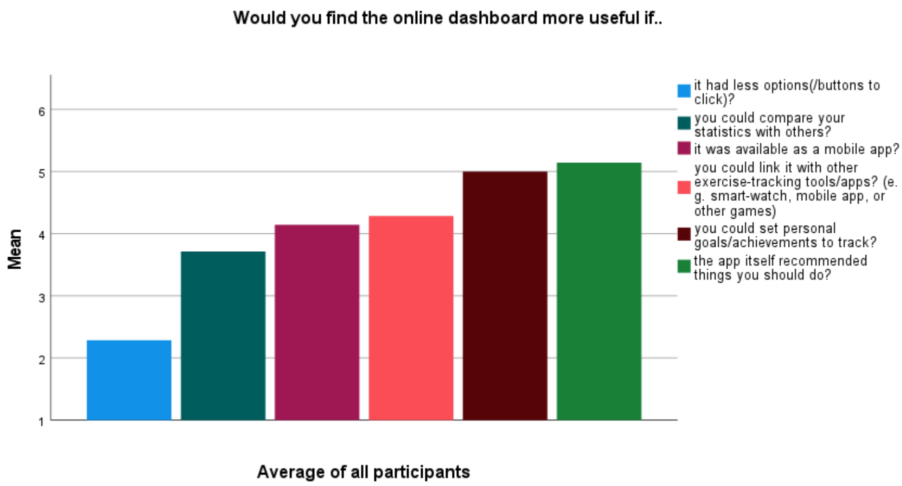


Figure 71: Graph: Rating whether potential changes to the online dashboard would make it less useful (1) or more useful (6)

The main points of improvements indicated by participants were in regards to setting and receiving goals to accomplish. The two right-most columns in the graph above (fig. 71) is in regards to this aspect. The highest-rated was that they would have liked if the application itself recommended things they should do. The second highest-rated was concerning the ability to set personal goals to accomplish within the application.

Participants also indicated that linking the application with other exercise-tracking tools would be a welcoming addition, and that the dashboard could be better suited for mobile. There were also a few comments on potential improvements for the dashboard:

The boxes (stat cards) below that contains info about calories burned, total playtime and so on, could be categorized a little better for each game. Example: If I would like to look at my Drone shooting stats, it would be nice to have a title and the stats below representing that game. A bit more clear than having to read "... in x game".

- Participant 1

In case of VR statistics in the app, it is not required to share this data with a server and simply store the data in local files.

- Participant 5

> *More data. Maybe ways of aggregating data. A way of settings personal achievements in the game maybe. I.e "Play every Monday four weeks straight".*
>
> *- Participant 5*

Participant 5 commented that it would be nice to have this data stored locally in the VR application. I think this is a data privacy concern, as the data logs are stored on an external server. Further, the comments generally point towards that more categorizations and aggregations of data collected would be interesting to have.

5.2.4 Users' Evaluation of the Overall Platform

Use in Daily Life

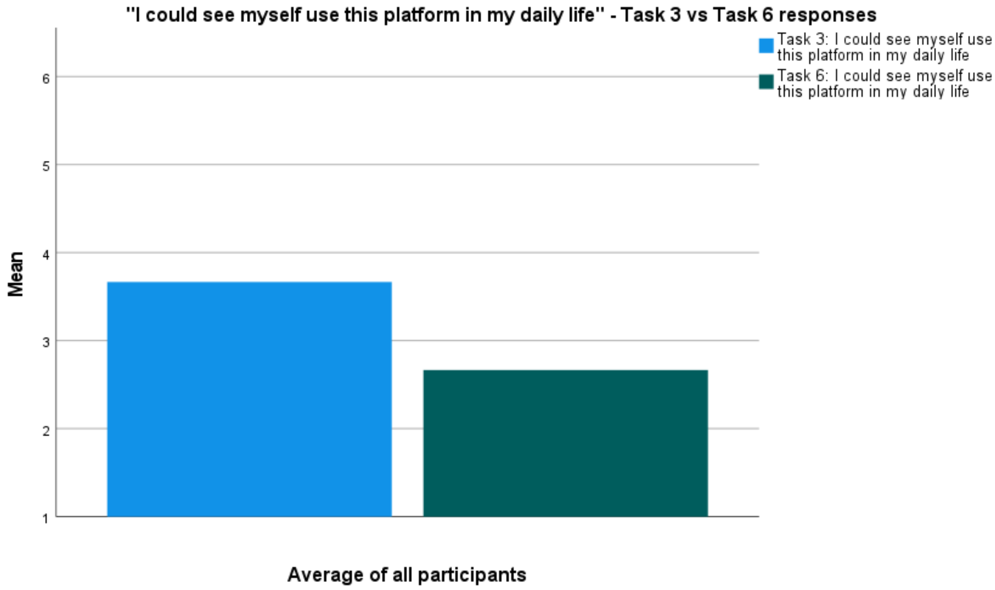


Figure 72: Graph: "I could see myself use this platform in my daily life" (Task 3 & Task 6 responses)

Figure 72 displays the combined rating from participants on whether they could see themselves use this platform in their daily life. This question was asked early in task 3 of the study, as well as in the last task of the study. Similar to the short-term versus long-term entertainment value graph (see fig. 68), it indicates that users initially had a higher spark of engagement or motivation to interacting with the platform. Over time, this engagement seems to have faltered. This could be to several reasons, but I think that one of the main issues is that the minigames lack the depth to sustain engagement in the long term. Lack of depth was quite prominent in the comments from the subjects as well. In addition, it is likely that many were more engaged early on as they were experiencing the application for the first time.

Participant 4 also commented with respect to this, which sums it up quite well:

> *For me to use this specific application in real life it would require more content. The content already implemented is fine, however, I think that for me to consider it an alternative for exercise I would need more games with more depth, as the current games can get a*

bit repetitive. For people who has played a lot of VR games, the initial "wow factor" of entering into VR is less present, and it becomes more important to provide more content and variation. One issue with most VR games, in my opinion, is that they lack depth. In the context of this project, that is of course not expected, but would be interesting to address.

- Participant 4

Completing tasks remotely

A substantial goal of this two-week study experiment was to evaluate whether the *ExerIsland* platform fulfilled its purpose in enabling the possibility of running *remote user studies*. Directly relevant to *focus A: Technologies to enable remote user studies*, and *research question 1: How to develop an exergame platform that enables the ability to run remote user studies?*

One of the larger concerns was how study tasks could be distributed and explained to participants. Additionally, the subjects would lack some of the potential direct 1-on-1 facilitation by me as an experiment facilitator. Though, I did offer some facilitation remotely if they would want it.

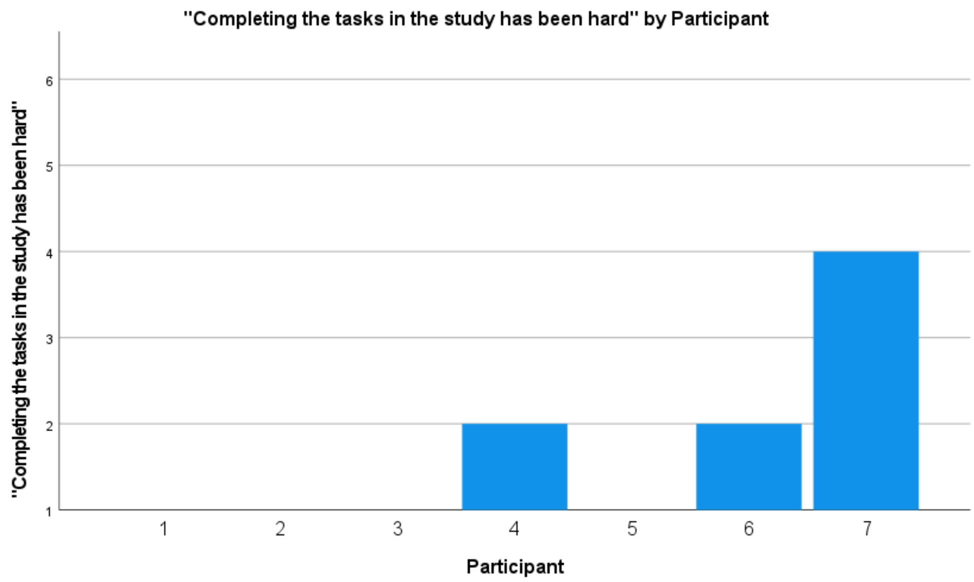


Figure 73: Graph: "Completing the tasks in this study has been hard"

Participants rated whether they found the tasks in the study hard to complete (fig. 73 above) on a Likert scale from 1 (not true) to 6 (very true). These results are very positively aligned. One participant rated the difficulty to a 4, while everyone else rated it to 2 or lower. This means that the explanations through task sheets (explained in section 4.2.3) combined with the usability of the overall platform made it intuitive for users to complete tasks.

Despite these promising results, I must add that there were some participants with no reported game logs from week two (described in section 5.2.1 above). This raises a few concerns:

Firstly, with this remote user study, there was no way for me to ensure that participants did not have any issues on their client-side when using the VR application. For example, if they were not connected to the internet, the logs from VR gameplay would not be uploaded to the backend. There should potentially have been some way to notify users in the VR app in case this was an issue.

Secondly, due to the nature of anonymous participation in a remote user study, I can not tell which people are struggling with completing tasks or not following up. It could potentially be that they had issues with the application, were misunderstanding assignments, or just decided not to follow up on tasks. I tried to mitigate this by telling subjects to contact me if there were any concerns, but then it would still be the participant's responsibility to reach out to me if they would need it.

Lastly, I must add that the second week of the study did put less pressure on participants performing tasks in the VR application. This means that the lack of involvement within the VR application in the second week could be an experimental design issue from my end. Most tasks in regards to interaction with the VR application was *optional* in the second week. I thought it could more easily display if participants were motivated to boot up the games because of the intervention (online dashboard) rather than asking them to do specific tasks within VR. Nevertheless, there were some flaws in the design of the two-week experiment. It is potentially part of why some participants did not have any logs from the second week of the study if they only did obligatory tasks.

In a remote user study like this, there will always be some layer of trust towards participants. As a researcher, I had to trust that subjects took the tasks of the user study seriously enough to complete them as described. Although the data in this remote two-week experiment does display some issues, I am overall very satisfied with the effort participants put into this study.

6 Discussion

In this chapter, I contextualize and interpret the results from the experiments and work conducted concerning the focus areas defined for the thesis. The discussion for *Focus A: Technologies to enable remote user studies* revolves around evaluating the ExerIsland platform's capabilities in enabling remote user studies. *Focus B: Gamification techniques for exercise motivation*, on the other hand, is a discussion around the relevant results from the experiments that I conducted.

6.1 Focus A: Technologies to enable remote user studies

For Focus A, there is mainly one research question I aim to provide relevant answers to:

Research Question 1: How to develop an exergame platform that enables the ability to run remote user studies?

The main relevant insights into this research question are gathered through evaluating the ExerIsland platform I developed for this specific purpose.

6.1.1 Evaluation of the ExerIsland Platform

I have tried to categorize some of the main evaluation points for the developed platform below. In each of these sections, I try my best to give my interpretation of the data collected throughout the study. I also try to put out my own personal perspective on how the system worked from the standpoint of a researcher conducting a remote user study.

Remote Anonymous Participation & Linking Data

In the ExerIsland Platform, new users entering the study received a generated *short-ID* which was 3-character long and only digits. This ID was stored in the backend and linked to their individual collected application data. Participants were told to write or remember this ID and use it for answering research questionnaires throughout study tasks.

First of all, I want to mention that this system did in practice not have any issues with the present user study. All participants understood this concept correctly and were able to provide the correct ID when answering surveys. I think a benefit of having such a simple system for remote anonymous participation in smaller studies is that it makes it easy for the participants to remember their personal ID. In this case, it was even more important, as participants were actively interacting in VR. I wanted to distribute a simple ID for them to remember as soon as they were outside of the VR setting.

With all the data from participants being linked to a simple 3-digit ID, it made it very simple to synthesize data from different sources together. In other words; when combining all the data points from surveys and logs, I could quickly combine these, as all had a common unique identifier with the participant's ID.

In the perspective of a small-sized user study, it made sense to do it this way. Although the short-ID being restricted to 3 digits means that it is capped at 1000 participants. Further, the 3-digit ID is what users use to identify themselves when logging data to the backend. It is also used as their login information (username and password) in the online dashboard application.

This poses a few security threats; Firstly, participants could very simply spoof the short-ID of someone else and upload logs that would seem to be from another user. This could be hard to recognize as a researcher, and could potentially damage the overall outcome of the user study. Secondly, malicious users could potentially do a very simple brute-force attack in order to get access and see the information and logs of a different user in the platform.

It was not considered within the scope of this thesis to cover these specific security issues of the platform. The main reasons were that the user study was on a relatively small scale (7 participants). Additionally, it did not include any sensitive information linked to users in the platform.

However, I will say that this is a rather substantial security flaw of the ExerIsland platform. For any larger remote user studies that seek to use some anonymous identification method, I would recommend paying a lot of attention to this system. Especially if users have any sensitive data linked to them.

In my opinion, one of the simplest and better ways of solving this issue would be to use a public-private key-pairing technique. The user could generate a private key on the client-side and distribute their public key to the backend. Then, the *observer/researcher* could receive access to the data from this specific public key on the backend. This would make it much harder to spoof other users, and it would also provide an immediate way to encrypt data in the database.

Distributing Tasks & Following up Participants

When it came to the distribution of tasks in the study, this was not an implemented part of the ExerIsland platform. The applications were developed with ease-of-use in mind, focusing on making functionality easy to understand from a user perspective. The tasks were distributed through task sheets describing exactly what participants ought to do for specific parts of the study. This method deemed quite successful. It seems like the active use of screenshots and detailed explanations worked well enough for participants to understand and complete tasks.

However, the results in section 5.2.1 describe that some users did not actively engage in the VR application in the second week of the study. Yet, these participants reported in questionnaires that they were engaging with the VR application. As all participation was anonymous, it was not possible as a researcher to know which subjects were struggling or falsely reporting behaviour that did not actually take place. The ExerIsland system was lacking in terms of features for following up participants within the platform.

The system could greatly benefit from having a way for *observers* to communicate with the participants in a study. It could simply be a messaging or note system within the platform. This would enable a way for researchers to give information to participants while not uncovering their identity.

In fact, a built-in messaging system would have more benefits to the system. If this was added, it could also be used as a task-distribution system. If participants could retrieve messages on the client-side it could simplify this process of distributing tasks a lot. In my study, I sent out task sheets to every participant on different messaging platforms, which was a rather cumbersome process even for just *7* participants. Even worse is that this solution is not scalable at all. It gets worse the more subjects you have in a study. For example; if the study had 100 participants I would have to manually send task sheets to all of them.

I would suggest that remote user studies, especially at a larger scale, should include a task distribution system within the platform. In the specific case of a VR exergame study like this one, I would probably try to integrate this system into the VR application. This way, users could directly have a schedule of tasks that will happen, and through starting their application they would be presented with the current assignment. It would make the distribution of tasks much smoother. It would also potentially make the process easier for participants, as they could read the tasks within the application (VR hub in this case) they are using.

Logs from VR as Experiment Data

Some of the data points collected through activity logging in VR was displayed in section 5.2.1. A benefit of conducting user studies with VR is that you get easy access to many valuable types of data to investigate. With using a framework like *SteamVR* in Unity, data from all VR headsets are interpreted very similarly in the game engine.

The movement data from these sensors can be further aggregated. In turn, it could give interesting logs to investigate in the context of a user study, but also for individual participants to look at. With the currently proposed system created for logging (explained in section 3.4.6), it enabled the ability to quickly create and implement new `MetricLoggers` to log data relevant for the study.

To give an example; if someone would use this system to conduct a rehabilitation user study on people with physical arm deficits, one could quickly create and integrate a logger that would track arm movement of the different arms. Then, researchers could observe how arm movement could change over time or its effect in different scenarios.

However, I must add that the data-logs collected through the ExerIsland application is restricted to quite simple data types as it currently is. The system was developed for storing mostly *integers* in the database. It tracked metrics such as arm movement, playtime or other game-relevant metrics. Comparing ExerIsland with other proposed systems from the literature review, it is somewhat limited in what can be done with the data. For example, there were several examples of studies that logged data so precisely that *observers* could see a graphical replay of the activity performed by the participant. [19, 20, 23]

In context of the ExerIsland platform and with logging inside VR, I think it would be reasonably feasible to implement similar tracking and replay functionality. However, it would depend on the use-case if there would be benefits of such advanced data representation. In the case of the two-week remote user study conducted in this thesis; I would argue that a simpler data type was better. The first reason is that it simply is a much less time-consuming logging solution to go for in a remote VR user study platform. The second reason is that a simpler data type can be more valuable in a study where I want to do quick and objective comparisons of users with quantitative data. Although this will depend on the specific goal of the user study, as previously mentioned.

Strengths & Flaws of the ExerIsland Platform

The ExerIsland platform in itself and the implementation details for it (described in chapter 3) is my direct contribution in trying to give an answer to *RQ1: "How to develop an exergame platform that enables the ability to run remote user studies?"*. The solution I have created with the ExerIsland platform displays how a remote user study platform can be created. However, my work only describes one possible way of solving this problem. I developed it with my interpretations of how such a system should be built and with the context of investigating the problem area of *focus B: Gamification techniques for exercise motivation*.

I have listed my interpretation of the strengths and flaws of my proposed platform below. This gives a brief overview of potential good practices to follow and pitfalls to avoid when creating remote user study systems. It is based on my experiences and results from running the two-week study experiment with the ExerIsland platform.

Strengths

- Valuable tools to inspect data (on frontend) for researcher and participants.
- Generating participant user on backend (described in section 3.6.1) enables a seamless experience for participants joining a study.
- Short participant ID's makes it easy for subjects to remember after leaving the user study platform (VR app in this case).
- The proposed data models containing *Observers* and *Participants* enables a scalable platform where several user studies could be going on at once.
- The data models (observers having a list of accessible participants) and authorization layers in the backend made sure it was not possible for users to get unauthorized access to objects in the database.
- The proposed logging data models *Session*, *Activities* and *Metrics* are simple to understand. The design of these datamodels allows for quick expansion of different metrics to log from activities and access in the frontend application.
- Restricting *Metrics* to one simple data type (integer in this case), was efficient for quick quantitative data analysis on the retrieved logs. Although, it will depend on the specific user study whether more advanced data types will be required.
- *Timestamps* on all session logs. Having a timestamp on all session-objects was relevant for data analysis as I could create a precise timeline of events by participants.
- Continuous updates of logs from the client-side (see section 3.4.6). If the client would have some issues on their end (e.g. quitting the app forcefully or losing internet connection), logs would still be uploaded if everything was functional last time an activity was completed.
- The ExerIsland platform is structured and designed in such a way that expanding with new minigames for user studies is simple. New minigames can quickly be added by deriving from existing classes (see section 3.4.3) and hooking up to existing `MetricLoggers`.

Flaws

- Using the 3-digit short-ID for authorization is too simple and imposes a large security threat to the system.
- With the *GET /generateNewUser* not having any type of authentication from the client-side, this becomes a vulnerable entry-point that could be prone to a spamming attack on this exact entry-point.
- ExerIsland did not include any way to communicate directly to participants within the platform, this makes it very hard to follow up subjects struggling as they are anonymous.
- Task distribution should have been possible to do within the ExerIsland platform. This would present a more scalable solution for a larger study. It could also potentially be a better experience for participants as they could receive tasks in-app.
- ExerIsland was limited in that it included a hardcoded observerID which received participants when they registered. A more scalable solution would be to make users include a *study-ID* from the client side when they called to register for a study. In turn, this ID could link to the responsible *observer* on the backend.
- The system is currently developed with VR as the only user study application. Exploring ways to log and include data from other day-to-day activities (e.g. from a mobile app) could greatly improve the overall use-cases of the system.
- The games in the ExerIsland VR application were lacking in depth, which caused the long-term entertainment value to be too poor. I think this damaged the outcome of the specific two-week user study I did. More advanced and in-depth games with longer entertainment value could potentially provide better and more accurate results.

6.2 Focus B: Gamification techniques for exercise motivation

For *Focus B* there are two research questions I delved deeper into. I have described relevant insights and my interpretations of the results with respect to *research question 2* and *research question 3* below.

6.2.1 RQ2: How can exercise trackers motivate people to engage more in physical activity?

Concerning research question 2, both data from the questionnaire and practical two-week experiment is relevant to keep in mind.

Making Exercise Fun

In the results from the research questionnaire, it displayed that the group that used exercise trackers had a higher count of workout hours on average per week (see section 5.1.1). Further, there were a lot more people in the group that used exercise trackers which deemed exercise as *fun* (in fig. 55).

The results indicate a quite strong statistical significant different mean value in these variables. However, it only displays what I would say is a correlation. It is not enough to prove a causal relationship between the independent variable (exercise tracker) and the dependent variables ("*exercise is fun*" and *workout-hours per week*). From this simple questionnaire data it is not possible to verify whether people found exercise more fun *because* they were using an exercise tracker. It might as well be that they intrinsically enjoy exercise more, and thus spend more time exercising and using exercise tracker to see their progress.

From the two-week user study, the results were quite mixed. I found that some users were more active in the second week after the online dashboard was introduced, but most users were most active in the first week. I think that the relatively low long-term entertainment value (rated by participants, see fig. 68) damaged the outcome of my specific user study in this regard. The application was mostly engaging in the beginning, which caused a considerable spike in activity in the first week when the study started. I think participants were not motivated enough solely by the online dashboard to engage much more in the activities they had already tested out in the first week.

Even though it did not display increased engagement so well by the logs from the VR activities, participants did subjectively report increased engagement with the VR activities after being presented with the dashboard (see fig. 70). This further confirms my statement that despite users feeling increased engagement, they were simply too bored of the VR games to play them more. In comparison with the papers I explored with my literature study for *focus B* (section 2.3.3), my findings would go under the *'Motivation increased then fading'* category in terms of motivational results. Though, many of the studies from my literature review did not conduct experiments over a long enough period to verify whether their initial engagement was just a spark of motivation that would fade over time. In that regard, I am satisfied with having run a long enough study experiment to uncover that the initial spark of engagement did seem to fade over time in my specific user study.

I think my findings are valuable in terms of pointing towards some directions. There is some correlation between the use of exercise trackers and people exercising more while also having more *fun*. However, based on the research I conducted for my thesis, my findings are restricted to only indicating this correlation. Though, I do think that more randomized controlled trial experiments with exercise trackers could give more precise and intriguing findings. I consider this as part of the potential future directions within this research area (see section 7.4).

Self-Assessment & Drive to Improve

Findings in the research questionnaire indicated that users with experience with exercise trackers found personal statistics to be the most crucial aspect. Workout performance graphs and tracking personal goals was also rated very high.

I think this indicates that people are driven either by doing some self-assessment on their workouts or straight up have an intrinsic motivation to improve. With an exercise tracker, they can have specified concrete results that are very straightforward to use for their own evaluation. In turn, this enables their ability to analyze whether and how effective their recent exercises have been.

The findings from the two-week user study align well with this indication. Several participants reported an increased motivation after seeing their online dashboard. As I see it, exercise trackers seem to have at least two potential effects on people: Firstly, it can spark engagement in the users, making them want to go out and perform the said activity. Secondly, it can make the user want to push their limits further in the specific activity because they want to see a concrete improvement from their previous performances.

6.2.2 RQ3: How should data from games be presented to engage users more in activities?

In relation to research question 3, the data from participants evaluating the usefulness of features in the online dashboard is highly relevant. Further, data from the online questionnaire survey also gives pointers to how people prefer to have data presented to them. In this specific case, the online dashboard was a type of exercise tracker, but for exergames. Thus, I will evaluate and discuss results from both experiments concerning *research question 3*.

Total Statistics

In the questionnaire study and two-week user study, participants rated *general statistics* to have a high value of the different features in exercise trackers. I think that there is some psychological aspect to this for many people. I believe people are driven by seeing that the work they put in is meaningful in the bigger picture.

For example, there was a participant in the ExerIsland study that indicated amusement when realizing that he had moved his arms 2.5 kilometers total by looking at the combined statistics on the online dashboard. I think these overall statistics give people a chance to zoom out and look at the bigger picture. It also enables the ability for people to set certain milestones to accomplish, for example, that they want to reach a certain amount of calories burned within the end of the month or similar.

Categorization & Aggregation of Data

For the total statistics explained above, several participants indicated that they would like to have more categorizations and ways to aggregate this data.

If an exergame platform presents several different categories of activities, it should be possible for people to categorize and synthesize more interesting data together within the application. The ExerIsland's online dashboard was limited in having a set amount of *Statistic Cards* that it always displayed, and it could have been improved by having a wider set of varied types and aggregations of these simple statistics.

Furthermore, participants in both present studies did indicate that they evaluated the ability to track personal goals quite high. I think that visual display of progress towards goals is a very motivating factor for many people. It gives a clear mission for the individual to work towards, and all work put in can be directly visualized.

Graphs

The importance of graphs in exercise trackers were rated as relatively significant in the questionnaire study and the ExerIsland user study. I think graphs that display some measured metric over time is a very concrete tool to see progress.

With the ExerIsland platform, I realized that there are many different categorizations of time-frames that can be essential to include for these performance graphs. In exercise tracking applications, I think it should be a possibility for users to adjust this time-frame to what they seem as important. The ability for users to zoom in on a narrow timeline (e.g. day-to-day) or zooming out for example on a month-to-month basis can be very valuable.

From a development perspective, it is a relatively simple addition to include in an exercise tracker. Still, providing a very effective personal data evaluation tool. It enables an easy way for people to evaluate their own performance over a specific time-frame.

Session-Specific Data

An extra part I found fascinating from the user study with the ExerIsland platform was that users rated the *Session list* feature the most valuable in the online dashboard (see fig. 69). I thought that the metric graphs or statistic cards would be the highest rated features.

However, users seem to like having the ability to see a concrete overview over all their in-game sessions. This is the data in its purest form, displaying exactly what is stored in the *Session, Activity & Metric* objects in the database. It might be that users found this very valuable because it allowed them to inspect all data from the specific session at the same time.

I think that this session list with session-specific data could be seen as a *timeline* of logs. It allows users to inspect their specific statistics at any point where they interacted with the exergame platform.

7 Summary & Conclusions

7.1 Focus A: Technologies to enable remote user studies

7.1.1 Research Question 1

With respect to my first research question, "*How to develop an exergame platform that enables the ability to run remote user studies?*", I explored this topic through practical implementation of my remote user study platform, *ExerIsland*. In many ways, the technical details described in chapter 3 is me trying to give an answer to this question with my interpretation of how such a system should be made.

My *two-week user study* with the platform was quite successful. Participants continued through the whole process with performing tasks. Data from in-game activities were properly logged to the database, and the data collected was valuable for data analysis.

However, I will add that my research was limited to testing out the user study with a VR application and only 7 participants. As I pointed out in section 6.1 there are several flaws of the platform which could see improvements. The most present issues that are not solved with the ExerIsland platform is some facilitation and security features. For facilitation, there is no communication between the researcher and participant within the platform. The lack of this feature was an obvious flaw as it deemed hard to follow up whether subjects were completing tasks or struggling in a remote and anonymous experiment setting. There is also no internal way to distribute tasks. For security features, the most substantial issue is that the *Short-ID* mechanism created for participants is too simple and imposes several security threats (e.g. spoofing).

Although my solution has its flaws, I would say that *ExerIsland* can be seen as a possible way to create a remote user study system. I think my experiences and technical details provided can be relevant for others looking to build similar platforms. My solution can serve as an inspiration for others, but they have to take their specific user study context in mind when designing their system.

7.2 Focus B: Gamification techniques for exercise motivation

7.2.1 Research Question 2

Concerning my second research question, "*How can exercise trackers motivate people to engage more in physical activity?*", I explored this topic through a practical two-week experiment with ExerIsland and an online questionnaire regarding exercise trackers.

My findings indicated a correlation displaying that people using exercise trackers seem to evaluate exercise as *fun* more often, compared to those that are not using trackers. A second correlation was also found, displaying that people using exercise trackers reported more workout-hours per week on average. The last main indication from my findings is that *personal statistics* is the most essential feature in these applications.

It could mean that the simple gamification of tracking exercise statistics and displaying this is enough to change people's mental attitude towards workouts. They could perceive the activity as more fun, which in turn could lead to increased activity levels.

Nevertheless, my findings are not enough to prove a causal relationship between these variables. My findings only serve as an indication of the potential benefits given by the gamifications with trackers. I think that several of my indications from this study should be further researched to verify if my indications are correct or not.

7.2.2 Research Question 3

In regards to my third research question, "*How should data from games be presented to engage users more in activities?*", I also explored this through the two-week experiment with ExerIsland and the online questionnaire.

From my experiments, I discovered that there were several ways individuals preferred to see their data. Some users fancied total statistics to look at, giving them certain personal milestones to aim for. While others favoured advanced categorizations of data and the ability to track their progress towards a specific goal. Other ways of displaying data reported relevant from users were visual graphs or the raw data from individual session logs.

In other words, my findings were quite varied, and I would say that I am not able to provide a direct answer to this research question. Simply because I think there is not a defined *'correct'* way of displaying data to engage users. The answer to this question will be based on the personal preference of the person you are asking, and additionally, it will depend on the context of the game activity.

The existing literature also displayed diverse solutions on how to display data in their gamified activities. It aligns well with the varied findings from my experiments. People clearly have divided opinions on how they prefer to display their data, and data display will depend on the context. Simply put, some data could be meaningful to present in overall statistics, while other data is more meaningful to display graphically.

7.3 Limitations

My main limitation regarding the two-week user study with the ExerIsland platform was that I had a limited amount of participants, with only *7*. Moreover, the fact that two of these users did not partake in any VR activities in the second week is a considerable weakness of the study. Due to the low amount of participants, I had to deviate from my original study design that included a control group that never received access to the online dashboard. A study with more participants would likely increase the quality of the data from the experiment a lot. It would also justify keeping the *control group & treatment group* design of the study.

Another limitation is that the results of the two experiments conducted in this study are highly based on subjectively reported opinions. When people answer questionnaires, they might subconsciously try to guess what the objective of the experiment is, which could yield biased results. Further, the questionnaire responses are only a *snapshot* of the current subjective thoughts of the participants. Essentially, it does not necessarily capture the bigger picture or how peoples opinions could change over time.

Lastly, I would like to mention that there is some selection bias to be aware of in both my conducted experiments. My study with the *ExerIsland* platform required users to have personal access to VR. These are people with a special interest in game technologies, and does not necessarily represent the opinions of a wider audience that well.

For the online questionnaire I distributed, the selection bias is not as concrete. The issue with distributing a questionnaire online is that there exist some self-selection bias. Basically, it is more likely that someone with a higher interest in exercise or exercise trackers chose to respond to the questionnaire because they found it interesting. In turn, the data here does not necessarily capture the opinions of a wider unbiased audience.

7.4 Future Directions

I found the research field within *remote user studies* to be quite limited. More studies, similar to mine, with practical implementation and experiment, could be beneficial. Particularly because remote user studies have several benefits. It can enable large-scale data collection from a wide audience in the world. It is also safe to conduct in the middle of a pandemic restricting physical interaction.

More randomized controlled trial experiments with a practical use of exercise trackers could yield more precise findings. I think more studies should investigate whether people find *physical activity* to be more *fun* with the addition of exercise trackers.

I found the long-term studies regarding gamifications to engage people more with physical activity to be quite limited. I think there should be a greater focus on the effects on people's exercise tendencies over more prolonged periods. This could display if it affects peoples intrinsic motivation when exposed to the gamification over a specific time period.

The existing literature is very rich in content for medical or rehabilitation purposes, and these fields are of high importance. Yet, I think the research field would benefit a lot from having more research concerning how we generally can increase the physical activity levels in our relatively *sedentary* society. With an increased physical activity level, we would likely have a reduced need for physical rehabilitation. In turn, this could increase life quality and lifespan of individuals, as well as reduce the potential economical burden on the society.

7.5 Concluding Thoughts

In regards to my first focus area, *Focus A: Technologies to enable remote user studies*, I am happy with my overall outcome in displaying a possible solution to enable remote user studies with *ExerIsland*. Even though my experiment with the platform was limited to only 7 participants, I feel that my results and experiences that I have shared in this thesis are fairly valuable. I think, and hope, that my study was able to shed some light on the rather unexplored topic of remote user studies.

Further, I find the topic of *Focus B: Gamification techniques for exercise motivation* to be very important in our current society. I have tried to explore specifically *exercise tracking* as the presented gamification. I am happy that my two-week experiment with the exergame platform was long enough to discover some of the issues with long-term engagement, as I found long-term studies to be somewhat limited in the literature. Due to these reasons, I think my study is quite unique and has a different angle than the existing literature. The insights collected from my practical experiments may not have resulted in direct answers but has shed light on potential crucial aspects to consider in future directions within this field of research.

Bibliography

[1] Unity - manual: Coroutines. https://docs.unity3d.com/Manual/Coroutines.html. (Accessed on 12/05/2021).

[2] flow – apa dictionary of psychology. https://dictionary.apa.org/flow. (Accessed on 11/05/2021).

[3] Conference, I. H. 2002. Constitution of the world health organization. 1946. *Bulletin of the World Health Organization*, 80(12), 983 – 984.

[4] Caspersen, C. J., Powell, K. E., & Christenson, G. M. 1985. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public health reports*, 100(2), 126.

[5] Kohl, H. W., Craig, C. L., Lambert, E. V., Inoue, S., Alkandari, J. R., Leetongin, G., & Kahlmeier, S. 2012. The pandemic of physical inactivity: global action for public health. *The Lancet*, 380(9838), 294–305. URL: https://www.sciencedirect.com/science/article/pii/S0140673612608988, doi:https://doi.org/10.1016/S0140-6736(12)60898-8.

[6] Andersen, L. B., Mota, J., & Di Pietro, L. 2016. Update on the global pandemic of physical inactivity. *Lancet (London, England)*, 388(10051), 1255–1256.

[7] Organization, W. H. Noncommunicable diseases. https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases. (Accessed on 21/04/2021).

[8] Lakicevic, N., Gentile, A., Mehrabi, S., Cassar, S., Parker, K., Roklicer, R., Bianco, A., & Drid, P. 2020. Make fitness fun: Could novelty be the key determinant for physical activity adherence? *Frontiers in Psychology*, 11.

[9] Deterding, S., Sicart, M., Nacke, L., O'Hara, K., & Dixon, D. 2011. Gamification. using game-design elements in non-gaming contexts. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '11, 2425–2428, New York, NY, USA. Association for Computing Machinery. URL: https://doi.org/10.1145/1979742.1979575, doi:10.1145/1979742.1979575.

[10] Deterding, S., Dixon, D., Khaled, R., & Nacke, L. 2011. From game design elements to gamefulness: defining" gamification". In *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments*, 9–15.

[11] Teixeira, P. J., Carraça, E. V., Markland, D., Silva, M. N., & Ryan, R. M. 2012. Exercise, physical activity, and self-determination theory: a systematic review. *International journal of behavioral nutrition and physical activity*, 9(1), 1–30.

[12] Sunnaas rehabilitation hospital - sunnaas sykehus. https://www.sunnaas.no/sunnaas-rehabilitation-hospital. (Accessed on 10/05/2021).

[13] Spill deg bedre - Spill deg bedre. https://www.spilldegbedre.no/. (Accessed on 10/05/2021).

[14] Psious | Virtual Reality Platform for Psychology and Mental Health. https://psious.com/. (Accessed on 10/05/2021).

[15] Peretti, A., Amenta, F., Tayebati, S. K., Nittari, G., & Mahdi, S. S. Jul 2017. Telerehabilitation: Review of the state-of-the-art and areas of application. *JMIR Rehabil Assist Technol*, 4(2), e7. URL: http://rehab.jmir.org/2017/2/e7/, doi:10.2196/rehab.7511.

[16] Koivisto, J. & Hamari, J. 2019. Gamification of physical activity: A systematic literature review of comparison studies. In *3rd International GamiFIN Conference, GamiFIN 2019*. CEUR-WS.

[17] Postolache, O., Hemanth, D. J., Alexandre, R., Gupta, D., Geman, O., & Khanna, A. 2021. Remote monitoring of physical rehabilitation of stroke patients using iot and virtual reality. *IEEE Journal on Selected Areas in Communications*, 39(2), 562–573. doi:10.1109/JSAC.2020.3020600.

[18] Choi, A., Noh, S., & Shin, H. 2020. Internet-based unobtrusive tele-monitoring system for sleep and respiration. *IEEE Access*, 8, 76700–76707.

[19] Howie, S. & Gilardi, M. 2020. Virtual observations: a software tool for contextual observation and assessment of user's actions in virtual reality. *Virtual Reality*, 1–14.

[20] Soufian, M., Nefti-Mezian, S., & Drake, J. 2020. Toward kinecting cognition by behaviour recognition-based deep learning and big data. *Universal Access in the Information Society*, 1–19.

[21] Lagun, D. & Agichtein, E. 2011. Viewser: Enabling large-scale remote user studies of web search examination and interaction. In *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '11, 365–374, New York, NY, USA. Association for Computing Machinery. URL: https://doi.org/10.1145/2009916.2009967, doi:10.1145/2009916.2009967.

[22] Andreasen, M. S., Nielsen, H. V., Schrøder, S. O., & Stage, J. 2007. What happened to remote usability testing? an empirical study of three methods. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, 1405–1414, New York, NY, USA. Association for Computing Machinery. URL: https://doi.org/10.1145/1240624.1240838, doi:10.1145/1240624.1240838.

[23] Tsekleves, E., Paraskevopoulos, I. T., Warland, A., & Kilbride, C. 2016. Development and preliminary evaluation of a novel low cost vr-based upper limb stroke rehabilitation platform using wii technology. *Disability and Rehabilitation: Assistive Technology*, 11(5), 413–422.

[24] Madary, M. & Metzinger, T. K. 2016. Real virtuality: a code of ethical conduct. recommendations for good scientific practice and the consumers of vr-technology. *Frontiers in Robotics and AI*, 3, 3.

[25] Nishiyama, Y., Ferreira, D., Eigen, Y., Sasaki, W., Okoshi, T., Nakazawa, J., Dey, A. K., & Sezaki, K. 2020. Ios crowd–sensing won't hurt a bit!: Aware framework and sustainable study guideline for ios platform. In *Distributed, Ambient and Pervasive Interactions*, Streitz, N. & Konomi, S., eds, 223–243, Cham. Springer International Publishing.

[26] Vitabile, S., Marks, M., Stojanovic, D., Pllana, S., Molina, J. M., Krzyszton, M., Sikora, A., Jarynowski, A., Hosseinpour, F., Jakobik, A., Stojnev Ilic, A., Respicio, A., Moldovan, D., Pop, C., & Salomie, I. *Medical Data Processing and Analysis for Remote Health and Activities Monitoring*, 186–220. Springer International Publishing, Cham, 2019. URL: https://doi.org/10.1007/978-3-030-16272-6_7, doi:10.1007/978-3-030-16272-6_7.

[27] Merilampi, S., Mulholland, K., Ihanakangas, V., Ojala, J., Valo, P., & Virkki, J. 2019. A smart chair physiotherapy exergame for fall prevention – user experience study. In *2019 IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH)*, 1–5. doi:10.1109/SeGAH.2019.8882482.

[28] Jaume-i-Capó, A., Martínez-Bueso, P., Moyà-Alcover, B., & Varona, J. 2014. Interactive rehabilitation system for improvement of balance therapies in people with cerebral palsy. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 22(2), 419–427. doi:10.1109/TNSRE.2013.2279155.

[29] Nagle, A., Riener, R., & Wolf, P. 2016. How would you like to be rewarded? relating the big-five personality traits with reward contingency in a cognitive training puzzle game. In *2016 IEEE International Conference on Serious Games and Applications for Health (SeGAH)*, 1–7. doi:10.1109/SeGAH.2016.7586281.

[30] Deacon, M., Parsons, J., Mathieson, S., & Davies, T. C. 2018. Can wii balance? evaluating a stepping game for older adults. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 26(9), 1783–1793. doi:10.1109/TNSRE.2018.2862146.

[31] S. V. Machado, F., Casagrande, W. D., Frizera, A., & da Rocha, F. E. M. 2019. Development of serious games for neurorehabilitation of children with attention-deficit/hyperactivity disorder through neurofeedback. In *2019 18th Brazilian Symposium on Computer Games and Digital Entertainment (SBGames)*, 91–97. doi:10.1109/SBGames.2019.00022.

[32] Bortone, I., Leonardis, D., Solazzi, M., Procopio, C., Crecchi, A., Bonfiglio, L., & Frisoli, A. 2017. Integration of serious games and wearable haptic interfaces for neuro rehabilitation of children with movement disorders: A feasibility study. In *2017 International Conference on Rehabilitation Robotics (ICORR)*, 1094–1099. doi:10.1109/ICORR.2017.8009395.

[33] Pirovano, M., Mainetti, R., Baud-Bovy, G., Lanzi, P. L., & Borghese, N. A. 2016. Intelligent game engine for rehabilitation (iger). *IEEE Transactions on Computational Intelligence and AI in Games*, 8(1), 43–55. doi:10.1109/TCIAIG.2014.2368392.

[34] Reyes, R., Chua, J. E., & Tumibay, M. 2020. Gamification of balance exercises with intervention techniques and feedback system using intervention, motivation, and game design analyses: A pre-clinical study. In *2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS)*, 158–162. doi:10.1109/CBMS49503.2020.00037.

[35] Wilkinson, A., Tong, T., Zare, A., Kanik, M., & Chignell, M. 2018. Monitoring health status in long term care through the use of ambient technologies and serious games. *IEEE Journal of Biomedical and Health Informatics*, 22(6), 1807–1813. doi:10.1109/JBHI.2018.2864686.

[36] Lindberg, R., Seo, J., & Laine, T. H. 2016. Enhancing physical education with exergames and wearable technology. *IEEE Transactions on Learning Technologies*, 9(4), 328–341. doi:10.1109/TLT.2016.2556671.

[37] Mystakidis, S., Cachafeiro, E., & Hatzilygeroudis, I. 2019. Enter the serious e-scape room: A cost-effective serious game model for deep and meaningful e-learning. In *2019 10th International Conference on Information, Intelligence, Systems and Applications (IISA)*, 1–6. doi:10.1109/IISA.2019.8900673.

[38] Sureephong, P., Puritat, K., & Chernbumroong, S. 2016. Enhancing user performance and engagement through gamification: Case study of aqua republica. In *2016 10th International Conference on Software, Knowledge, Information Management Applications (SKIMA)*, 220–224. doi:10.1109/SKIMA.2016.7916223.

[39] Wollmann, T., Abtahi, F., Eghdam, A., Seoane, F., Lindecrantz, K., Haag, M., & Koch, S. 2016. User-centred design and usability evaluation of a heart rate variability biofeedback game. *IEEE Access*, 4, 5531–5539. doi:10.1109/ACCESS.2016.2601882.

[40] Kern, F., Winter, C., Gall, D., Käthner, I., Pauli, P., & Latoschik, M. E. 2019. Immersive virtual reality and gamification within procedurally generated environments to increase motivation during gait rehabilitation. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 500–509. doi:10.1109/VR.2019.8797828.

[41] Majid, M. S. H., Khairunizam, W., Ikram, K., Jing, L. M., Sahyudi, B. N., Zunaidi, I., Ariffin, M. A., Bakar, A. S., & Razlan, Z. M. 2018. Performance evaluation of a vr-based arm rehabilitation using movement sequence pattern. In *2018 IEEE 14th International Colloquium on Signal Processing Its Applications (CSPA)*, 123–128. doi:10.1109/CSPA.2018.8368698.

[42] Theriualt, A., Nagurka, M., & Johnson, M. J. 2014. Therapeutic potential of haptic theradrive: An affordable robot/computer system for motivating stroke rehabilitation. In *5th IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics*, 415–420. doi:10.1109/BIOROB.2014.6913812.

[43] Lu, F. & Lemonde, M. 2014. Reducing adolescent obesity with a social networking mobile fitness application. In *2014 IEEE 16th International Conference on e-Health Networking, Applications and Services (Healthcom)*, 429–434. doi:10.1109/HealthCom.2014.7001881.

[44] Goh, D. H.-L. & Razikin, K. 2015. Is gamification effective in motivating exercise? In *Human-Computer Interaction: Interaction Technologies*, Kurosu, M., ed, 608–617, Cham. Springer International Publishing.

[45] Kontadakis, G., Chasiouras, D., Proimaki, D., Halkiadakis, M., Fyntikaki, M., & Mania, K. 02 2020. Gamified platform for rehabilitation after total knee replacement surgery employing low cost and portable inertial measurement sensor node. *Multimedia Tools and Applications*, 79. doi:10.1007/s11042-018-6572-6.

[46] Seo, K., Kim, J., Ryu, H., & Jang, S. *'RehabMaster TM': A Pervasive Rehabilitation Platform for Stroke Patients and Their Caregivers*, 131–155. Springer London, London, 2014. URL: https://doi.org/10.1007/978-1-4471-6413-5_6, doi:10.1007/978-1-4471-6413-5_6.

[47] Gonçalves, A., Muñoz, J., Gouveia, E., Cameirão, M., & Bermúdez i Badia, S. 07 2019. Effects of prolonged multidimensional fitness training with exergames on the physical exertion levels of older adults. *The Visual Computer*. doi:10.1007/s00371-019-01736-0.

[48] Hocine, N., Gouaïch, A., & Cerri, S. A. 2014. Dynamic difficulty adaptation in serious games for motor rehabilitation. In *Games for Training, Education, Health and Sports*, Göbel, S. & Wiemeyer, J., eds, 115–128, Cham. Springer International Publishing.

[49] Su, W.-C., Yeh, S.-C., Lee, S.-H., & Huang, H.-C. 2015. A virtual reality lower-back pain rehabilitation approach: System design and user acceptance analysis. In *Universal Access in Human-Computer Interaction. Access to Learning, Health and Well-Being*, Antona, M. & Stephanidis, C., eds, 374–382, Cham. Springer International Publishing.

[50] Zuckerman, O. & Gal-Oz, A. 10 2014. Deconstructing gamification: Evaluating the effectiveness of continuous measurement, virtual rewards, and social comparison for promoting physical activity. *Personal and Ubiquitous Computing*, 18. doi:10.1007/s00779-014-0783-2.

[51] Osorio, G. 2018. Personal training mechanisms to encourage participation in interactive cycling. In *Smart Technology*, Torres Guerrero, F., Lozoya-Santos, J., Gonzalez Mendivil, E., Neira-Tovar, L., Ramírez Flores, P. G., & Martin-Gutierrez, J., eds, 189–198, Cham. Springer International Publishing.

[52] Zhao, Z., Arya, A., Orji, R., & Chan, G. 2020. Physical activity recommendation for exergame player modeling using machine learning approach. In *2020 IEEE 8th International Conference on Serious Games and Applications for Health (SeGAH)*, 1–9. doi:10.1109/SeGAH49190.2020.9201820.

[53] Segal, A. D., Lesak, M. C., Suttora, N. E., Silverman, A. K., & Petruska, A. J. 2020. irebot: An interactive rehabilitation robot with gesture control. In *2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, 5158–5161. doi:10.1109/EMBC44109.2020.9176696.

[54] Kordatos, G. & Modestos, S. 2019. Design and evaluation of a wearable system to increase adherence to rehabilitation programmes in acute cruciate ligament (cl) rupture. *Multimedia Tools and Applications*, 1–26.

[55] Aware – open-source context instrumentation framework for everyone. https://awareframework.com/. (Accessed on 07/05/2021).

[56] What are the harmful effects of virtual reality? - law technology today. https://www.lawtechnologytoday.org/2021/01/what-are-the-harmful-effects-of-virtual-reality/. (Accessed on 07/05/2021).

[57] Appelbaum, P. S., Roth, L. H., Lidz, C. W., Benson, P., & Winslade, W. 1987. False hopes and best data: consent to research and the therapeutic misconception. *The Hastings Center Report*, 17(2), 20–24.

[58] Vive™ | discover virtual reality beyond imagination. https://www.vive.com/us/. (Accessed on 30/05/2021).

[59] Oculus | vr-briller og utstyr. https://www.oculus.com/. (Accessed on 30/05/2021).

[60] firebase · pkg.go.dev. https://pkg.go.dev/firebase.google.com/go/v4. (Accessed on 08/05/2021).

[61] Flutter - beautiful native apps in record time. https://flutter.dev/. (Accessed on 08/05/2021).

[62] Csikszentmihalyi, M., Abuhamdeh, S., & Nakamura, J. *Flow*, 227–238. Springer Netherlands, Dordrecht, 2014. URL: https://doi.org/10.1007/978-94-017-9088-8_15, doi:10.1007/978-94-017-9088-8_15.

[63] Bartle, R. 1996. Hearts, clubs, diamonds, spades: Players who suit muds. *Journal of MUD research*, 1(1), 19.

[64] Nakamura, J. & Csikszentmihalyi, M. 2009. Flow theory and research. *Handbook of positive psychology*, 195–206.

[65] Introduction to widgets - flutter. https://flutter.dev/docs/development/ui/widgets-intro. (Accessed on 11/05/2021).

[66] .net - generating human-readable/usable, short but unique ids - stack overflow. https://stackoverflow.com/questions/9543715/generating-human-readable-usable-short-but-unique-ids. (Accessed on 12/05/2021).

[67] Independent samples t test - spss tutorials - libguides at kent state university. https://libguides.library.kent.edu/spss/independentttest. (Accessed on 16/05/2021).

A Literature Review Spreadsheets

These are the results of the literature search conducted for my background literature study.

A.1 Focus A Literature Spreadsheet

This spreadsheet can be seen here:

https://docs.google.com/spreadsheets/d/1Q6V3K3mmAlwlRuSdg0v-x7N4bqCgTGzQ14rkJEHUfvs/edit?usp=sharing

Title	Authors	Publish Ye	Publication	Study-data	Research approac	Hardware Technologies	Research topic(s)	Software Technologie	Remote context	Display of data	Metrics recorded	Main Findings	Other Findings	Link
Remote Monitoring of Physical Rehabilitation of Stroke Patients Using IoT and Virtual Reality	Octavian Postolache; D. Jude Hem	2021	Journal	Quantitative	R&D	Wearable-technology	Remote Monitoring, IoT, VR	Unity3D, Mobile App	Rehabilitation, Medical	Session graphs	Score, Left-Hand score, Right-Hand score	Can faciliate physicians and physical therapist about training outcome, useful to personalize exercises for patients. In turn this helps patients achieve better rehabilitation results in short period of time process.	Highly-interactive engagement.	https://ieeexplore.ieee.org/document/918398
Internet-Based Unobtrusive Tele-Monitoring System for Sleep and Respiration	Ahyoung Choi; Sangwoo Noh; Han	2020	Journal	Quantitative	R&D	Wearable-technology, Mobile, Computer	Remote Monitoring, IoT	Mobile App, C	Medical	Metric over time graphs	Respiration rate	System was useful for sleap health management. Can also be utilized to detect acute diseases such as systole or respiratory failure.	Potentially useful system to e.g. monitor elderly living alone, or monitor people living in places where it is difficult to provide medical services in person.	https://ieeexplore.ieee.org/document/907522
Virtual Observations: a software tool for contextual observation and assessment of user's actions in virtual reality	Scott Howie & Marco Gilardi	2020	Journal	Mixed	R&D	VR, Computer	Remote Monitoring, Remote User Studies, Logging, VR	Unity3D	Experiment, Observation	Graphical/Video	Position, time-stamps, interactions	Datasets recorded in Unity can be exported as XML and shared for others to display and review the activity through a simulation of the activities performed by the user.	Their Virtual Observation software offers detailed reconstruction of low-level information and visual information of user actions during simulations for both live and offline observations. Beneficial for researchers, developers and educators	https://link.springer.com/article/10.1007/s1005
Toward Kinecting cognition by behaviour recognition-based deep learning and big data	Majeed Soufian, Samia Nefti-Mezi	2020	Journal	Quantitative	R&D	Camera tracking, Wearable-technology	Remote Monitoring, IoT, Machine Learning, Logging	C	Medical, Observation	Graphical/Video	Position, time-stamps, interactions, movement speed	"By monitoring and recognizing important human behaviour such as eating, walking, furniture crawling, physical (in)activities, the system can calculate the main attributes related to key indicators of health events that are necessary for the diagnosis and prognosis of cognitive-related conditions in older adults and hence prolong independent living"	"While the results in this pilot study are based on a small number of elderly residents in a short period of time, they imply that within the MiiHome project it is possible to obtain indicators of cognitive decline. As a result, further comprehensive and longer studies with continuous cognitive assessment of many more participants are to be planned for full clinical validation of these indications."	https://link.springer.com/article/10.1007/s1020
ViewSer: enabling large-scale remote user studies of web search examination and interaction	Dmitry Lagun, Eugene Agichtein	2011	Proceedings	Quantitative	R&D	Computer, Camera tracking	Remote User Studies	Javascript, Search engine	Experiment, Observation	Graphical/Video	Position, Eye-tracking, Clicks	Software developed ViewSer: By restricting results from search-engine to viewport created by mouse-position, they could mimick "eye-tracking" quite accurately through this method. The results from remote user studies induced similar behaviour as in-house physical eye-tracking. In turn, this could enable large scale studies to be run with these "mimicked" eye-tracking software which enables large-scale web search behavioral evaluation studies for a much lower cost than in-lab studies.		https://dl.acm.org/doi/abs/10.1145/2009916.20
What happened to remote usability testing?: an empirical study of three methods	Morten Sieker Andreasen, Henrik V	2007	Proceedings	Mixed	Systematic experimental comparison of three methods for remote usability testing and a conventional laboratorybased think-aloud test	Computer	Remote User Studies	Virtual Network Computing	Experiment, Observation	Data tables	System Usability	Remote synchronous usability testing provided almost entirely similar results as conventional on-site method in the same amount of time. The asynchronous method however were more time-consuming and identified fewer usability problems.	Physical presence of research facilitator does not seem important. In fact their results indicate that some participants thought that the method without facilitator was less stressful. One participant stated: "I liked this test method better than the traditional method where the test leader looks over your shoulder."	https://dl.acm.org/doi/abs/10.1145/1240624.12
Development and preliminary evaluation of a novel low cost VR-based upper limb stroke rehabilitation platform using '	Emmanuel Tsekleves, Ioannis The	2014	Journal	Mixed	R&D	Wearable-technology, Videogame console, VR	Remote Monitoring, Rehabilitation	Three layers: data access a	Rehabilitation, Medical, (	Graphical/Video	Position, Arm-movement	Recorded data from user sessions could be visualized directly to the physiotherapist and seen as beneficial for providing the remote progress, assessment and interaction between therapist and patient.	Fused Nintendo Wii remote acceleration and gyroscope data together to track movement and get 6 degrees of freedom with the Wii remote alone.	https://www.tandfonline.com/doi/full/10.3109/1
Real Virtuality: A Code of Ethical Conduct. Recommendations for Good Scientific Practice and the Consumers of VF	Michael Madary and Thomas K. M	2016	Review Article	Literature Review	Literature Study		VR, Ethics, Online Ethics		Observation, Experiment, Industry			research with VR in an online study raises the likelihood that users may give consent for data collection in a particular virtual context but then become unaware of the continued data collection as the user changes context. I.e. make the it clear to the user that	principle of non-maleficence: do no harm. "No experiment should be conducted using virtual reality with the foreseeable consequence that it will cause serious or lasting harm to a subject." For informed consent: Inform	https://www.frontiersin.org/articles/10.3389/fro
IOS Crowd–Sensing Won't Hurt a Bit!: AWARE Framework and Sustainable Study Guideline for iOS Platform	Yuuki Nishiyama, Denzil Ferreira, Y	2020	Proceedings	Quantitative	R&D	Mobile, Mobile sensors	Remote User Studies	AWARE Framework	Experiment, Observation	Metric over time graphs	Mobile sensor data, Questionnaires, time-stamps	Created MCS framework for iOS, AWARE-iOS, based on AWARE framework for Android. Created guidelines for mobile crowd sensing studies with iOS. They tested their platform with three different ways of storing data: JSON, SQLITE and CSV. Each having different advantages over each other depending on context.	Created a Mobile App Study platform which lets users join studies with an ID, then lets them see the data they are sending and answer surveys within the platform.	https://link.springer.com/chapter/10.1007/978-
Medical Data Processing and Analysis for Remote Health and Activities Monitoring	Salvatore Vitabile,Michal Marks,Dr	2019	Book	Literature Review	Literature Study		Remote Monitoring		Observation, Medical, Industry			Chapter 4 gives an example of how a system architecture of a big data IoT remote monitoring system can look like in practice. Inspiration.		https://link.springer.com/chapter/10.1007/978-

A.2 Focus B Literature Spreadsheet

This spreadsheet can be seen here:

https://docs.google.com/spreadsheets/d/1oGQFIvWqajvStkuM9JoYW2k25QNP3rVaXBpjGxf0kh8/edit?usp=sharing

Title	Authors	Publish Ye	Publication	Study-data	Research approac	Technology used	Research topic	Participants age	Participants	Study length	Intervention	Motivation results	Long-term motivation resul	Other results
A Smart Chair Physiotherapy Exergame for Fall Prevention – User Experience Study	Sari Merilampi; Kyle Mulholland; Ve	2019	Proceedings	Quantitative	Survey	Mobile application	User experience	16-89	29	1 session	Exergame	Increased	No	Increased social interaction
Interactive Rehabilitation System for Improvement of Balance Therapies in People With Cerebral Palsy	Antoni Jaume-i-Capó; Pau Martíne	2014	Journal	Quantitative	Experiment	Camera body tracking	Task performance	20-65	9	24 weeks	Exergame, Personalization	Increased then fading	No	Increased task performance
How would you like to be rewarded? Relating the Big-Five personality traits with reward contingency in a cognitive training puzzle game	Aniket Nagle; Robert Riener; Peter	2016	Proceedings	Quantitative	Experiment	Computer	Task motivation	20-50	90	1 session	Reward-system, Personalization	Increased	No	None
Can Wii Balance? Evaluating a Stepping Game for Older Adults	Mark Deacon; John Parsons; Sear	2018	Journal	Mixed	Experiment	Feet trackpad	Task performance	50<	104	1 session	Exergame	Increased	No	None
Development of Serious Games for Neurorehabilitation of Children with Attention-Deficit/Hyperactivity Disorder through Neurofeedback	Fabiana S. V. Machado; Wagner D.	2019	Proceedings	Mixed	Experiment	Computer	Task performance	>18	3	1 session	Game	Mixed	No	None
Integration of serious games and wearable haptic interfaces for Neuro Rehabilitation of children with movement disorders: A feasibility study	Ilaria Bortone; Daniele Leonardis; N	2017	Proceedings	Mixed	Experiment	VR	Exercise motivation	08-30	12	1 session	VR Immersion	Increased	No	None
Intelligent Game Engine for Rehabilitation (IGER)	Michele Pirovano; Renato Mainett	2014	Journal	Mixed	Experiment	Camera body tracking, Feet trackpad	Usability	68-82	7	1 session	Performance Feedback, Personalization	Increased	No	None
Gamification of Balance Exercises with Intervention Techniques and Feedback System using Intervention, Motivation, and Game Design Analyses: A Pre-Clinical Stu	Rosula Reyes; Jedd Emille Chua; N	2020	Proceedings	Quantitative	Experiment	Feet trackpad	Exercise motivation	22-26	18	1 session	Exergame	Increased	No	None
Monitoring Health Status in Long Term Care Through the Use of Ambient Technologies and Serious Games	Andrea Wilkinson; Tiffany Tong; At	2018	Journal	Quantitative	Experiment	Tablet	Life quality	70-90	27	12 weeks	Personalization	Increased	3 months or longer	Increased well-being
Enhancing Physical Education with Exergames and Wearable Technology	Renny Lindberg; Jungryul Seo; Teer	2016	Journal	Mixed	Experiment	Wearable-technology	Task motivation	10	61	1 session	Exergame	Increased	No	Increased task performance
Enter the Serious E-scape Room: A Cost-Effective Serious Game Model for Deep and Meaningful E-learning	Stylianos Mystakidis; Enrique Cach	2019	Proceedings	Mixed	Experiment	VR	Task performance	14-16	148	1 session	VR Immersion	Increased	No	Increased task performance
Enhancing user performance and engagement through gamification: Case study of aqua republica	Pradorn Sureephong; Kitti Puritat; S	2016	Proceedings	Quantitative	Experiment	Computer	Task performance		30	1 session	Reward-system, Socialization, Competition	Increased	No	Increased task performance
User-Centred Design and Usability Evaluation of a Heart Rate Variability Biofeedback Game	Thomas Wollmann; Farhad Abtahi;	2016	Journal	Mixed	Experiment	Wearable-technology	Task performance	40-67	11	1 session	Exergame, Reward-system	Increased	No	Increased task performance
Immersive Virtual Reality and Gamification Within Procedurally Generated Environments to Increase Motivation During Gait Rehabilitation	Florian Kern; Carla Winter; Dominik	2019	Proceedings	Quantitative	Experiment	VR	Exercise motivation	19-30	36	1 session	VR Immersion, Autonomy	Increased	No	Increased well-being
Performance evaluation of a VR-based arm rehabilitation using movement sequence pattern	Mohd Saiful Hazam Majid; W. Khair	2018	Proceedings	Mixed	Case-study	Camera body tracking	Task performance	22	1	1 session	Exergame	Increased	No	Increased task performance
Therapeutic potential of haptic TheraDrive: An affordable robot/computer system for motivating stroke rehabilitation	Andrew Theriualt; Mark Nagurka; N	2014	Proceedings	Mixed	Experiment	Wearable-technology	Exercise motivation		6	1 session	Adjusting difficulty	None	No	None
Reducing adolescent obesity with a social networking mobile fitness application	Fletcher Lu; Manon Lemonde	2014	Proceedings	Quantitative	Experiment	Mobile application	Exercise motivation	14-15	35	8 weeks	Socialization	Increased	No	Reduced BMI for overweight
Is Gamification Effective in Motivating Exercise?	Dion Hoe-Lian Goh; Khasfariyati R.	2015	Proceedings	Mixed	Experiment	Mobile application	Exercise motivation	28	100	4 weeks	Reward-system, Socialization, Competition	Increased	No	Competition demotivating for some
Gamified platform for rehabilitation after total knee replacement surgery employing low cost and portable inertial measurement sensor node	Gregory Kontadakis, Dimitrios Cha	2018	Journal	Mixed	Experiment	Wearable-technology	Task performance	64-80	10	2 weeks	Exergame, Performance Feedback	Increased	No	Increased task performance
‹RehabMaster TM ›: A Pervasive Rehabilitation Platform for Stroke Patients and Their Caregivers	Kyoungwon Seo; Jieun Kim; Hokyo	2014	Book	Mixed	Experiment	Camera body tracking	Exercise motivation	50	16	2 weeks	Competition, Adjusting difficulty	Increased	No	None
Effects of prolonged multidimensional fitness training with exergames on the physical exertion levels of older adults	Afonso Rodrigues Gonçalves, Joh	2019	Journal	Quantitative	Experiment	Camera body tracking	Task performance	67	31	12 weeks	Exergame, Music	Increased	3 months or longer	Not improved task performance compared to traditional methodolo
Dynamic Difficulty Adaptation in Serious Games for Motor Rehabilitation	Nadia Hocine, Abdelkader Gouaïch	2014	Proceedings	Mixed	Experiment	Computer	Task motivation	38-73	7	2 weeks	Adjusting difficulty	Increased	No	None
A Virtual Reality Lower-Back Pain Rehabilitation Approach: System Design and User Acceptance Analysis	Wu-Chen Su, rShih-Ching Yeh, Si-H	2015	Proceedings	Quantitative	Experiment	Wearable-technology	Usability	65	20	6 weeks	Exergame, Performance Feedback	Increased	No	None
Deconstructing gamification: evaluating the effectiveness of continuous measurement, virtual rewards, and social comparison for promoting physical activity	Oren Zuckerman, Ayelet Gal-Oz	2014	Journal	Quantitative	Experiment	Mobile application	Exercise motivation	20-54	99	2 weeks	Socialization, Reward-system, Performance Feedback	Increased	No	No difference between system with socialization and rewards, compa
Personal Training Mechanisms to Encourage Participation in Interactive Cycling	Gume Osorio	2018	Proceedings	Quantitative	Experiment	Physical instrument	Exercise motivation	36	12	4 days	Performance Feedback, Music	Increased	No	Combination of game + feedback + music gave best results for enjoy
Physical Activity Recommendation for Exergame Player Modeling using Machine Learning Approach	Zhao Zhao; Ali Arya; Rita Orji; Gerr	2020	Proceedings	Qualitative	Experiment	Wearable-technology, Mobile applica	Exercise motivation, Usability	23-36	5	1 session	Personalization	None	No	Personality type only has limited influence on peoples preference tov
iRebot: An Interactive Rehabilitation Robot with Gesture Control	Ava D. Segal; Mark C. Lesak; Neil E	2020	Proceedings	Quantitative	Experiment	Wearable-technology, Physical Instru	Exercise motivation, Usability	21-39	11	1 session	Exergame	Increased	No	Increased task performance, Enhanced engagement compared to st
Design and evaluation of a wearable system to increase adherence to rehabilitation programmes in acute cruciate ligament (CL) rupture	George Kordatos & Modestos Sta	2019	Journal	Quantitative	Experiment	Wearable-technology	Exercise motivation, Usability	17-25	15	1 session	Exergame	Increased	No	Hedonic motivation that derives from the pleasure or satisfaction to

B Referenced Code: Complete Examples

B.1 VR Application

Code for VR Unity application can be reached here:
gitlab.com/akerholten/exerisland-vr-exergame-hub

B.2 Golang Backend

Code for Golang backend can be reached here:
gitlab.com/akerholten/vr-health-and-wellness-remote-monitoring/-/tree/master/backend

Note: Some of the code examples here still use method and variable names like *patients* instead of *participants*. This is because not all parts of the code-base have been updated to follow the new naming conventions after the platform was changed from a rehabilitation focus to a remote user-study platform focus.

B.3 Flutter Frontend

Code for Flutter frontend can be reached here:
gitlab.com/akerholten/vr-health-and-wellness-remote-monitoring/-/tree/master/frontendflutter

Note: Some of the code examples here still use method and variable names like *patients* instead of *participants*. This is because not all parts of the code-base have been updated to follow the new naming conventions after the platform was changed from a rehabilitation focus to a remote user-study platform focus.

observerDashboard.dart

This code displays some of the complexity issues faced with Flutter frontend applications. The nested widgets quickly become a pyramid of doom, which is a common issue in several frontend frameworks.

```
import 'package:flutter/material.dart';
import 'package:frontendflutter/src/handlers/observerHandler.dart';
import 'package:frontendflutter/src/model_classes/patient.dart';
import '../handlers/tools.dart';
import '../components/modal_AddNewPatient.dart';
import '../constants/route_names.dart';
import '../constants/constants.dart';

class ObserverDashboard extends StatefulWidget {
  @override
  _ObserverDashboardState createState() => _ObserverDashboardState();
}

class _ObserverDashboardState extends State<ObserverDashboard> {
  final scaffoldKey = GlobalKey<ScaffoldState>();
  int patientCount = 0;

  Patient newPatient = new Patient();
  List<Patient> patients;
  bool _loading = false;

```

```
List<String> columnTitles = [
  'Name',
  'Email',
  'Note',
  'Age',
  'Goals',
  'Recent activity'
];

void _fillWithTempData() {
  patients = new List<Patient>();
}

void _getAllParticipants() async {
  setState(() {
    _loading = true;
  });

  List<Patient> tempPatients = new List<Patient>();
  tempPatients = await ObserverHandler.getAllPatients();

  setState(() {
    _loading = false;
    // In case it returns null, we don't want the screen to continuosly
    // spam the backend each update for patients it can't retrieve
    if (tempPatients == null) {
      patients = new List<Patient>();
    } else {
      patients = tempPatients;
    }
  });
}

void _showAddNewPatientModal() {
  showDialog(
      context: context,
      builder: (BuildContext context) {
        return AlertDialog(
          content: AddNewPatientModal(
            onPatientAdded: () {
              _getAllParticipants();
            },
          ),
        );
      });
}

@override
Widget build(BuildContext context) {
  double tableItemWidth =
      (Constants.pageMaxWidth * 0.75) / columnTitles.length;
  double tableItemHeight = 70;

  ScrollController _controller = new ScrollController();

  // data cannot be null
  if (patients == null) {
    _fillWithTempData();
    // Asynchronously collecting all participants
    _getAllParticipants();
  }

  Widget tableHeader() {
    return Container(
      padding: EdgeInsets.only(left: 16, right: 16),
      child: Column(
        mainAxisAlignment: MainAxisAlignment.spaceBetween,
```

```
        children: [
          Row(
            mainAxisAlignment: MainAxisAlignment.spaceBetween,
            children: (columnTitles.map(
              (item) => Container(
                alignment: Alignment.center,
                height: tableItemHeight,
                width: tableItemWidth,
                child: SelectableText(
                  item,
                  style: Theme.of(context).textTheme.headline6,
                ),
              ),
            )).toList(),
          ),
          Container(
            height: 1,
            width: double.maxFinite,
            color: Theme.of(context).dividerColor,
          ),
        ],
      ),
    );
  }

  Widget tableRows() {
    return Container(
      padding: EdgeInsets.only(bottom: 16, left: 16, right: 16),
      child: Container(
        height: Constants.pageMaxHeight,
        child: ListView(
          physics: const AlwaysScrollableScrollPhysics(),
          controller: _controller,
          shrinkWrap: true,
          children: (patients
              .map((patient) => FlatButton(
                    onPressed: () => Navigator.of(context).pushNamed(
                        Routes.SpecificPersonDashboard,
                        arguments:
                            PatientDashboardArguments(patient.shortID)),
                    child: Column(
                      mainAxisAlignment: MainAxisAlignment.spaceBetween,
                      children: [
                        Row(
                          mainAxisAlignment: MainAxisAlignment.spaceBetween,
                          children: [
                            Container(
                              alignment: Alignment.center,
                              height: tableItemHeight,
                              width: tableItemWidth,
                              child: SelectableText(
                                  patient.firstName + " " + patient.lastName),
                            ),
                            Container(
                              alignment: Alignment.center,
                              height: tableItemHeight,
                              width: tableItemWidth,
                              child: SelectableText(patient.email),
                            ),
                            Container(
                              alignment: Alignment.center,
                              height: tableItemHeight,
                              width: tableItemWidth,
                              child: SelectableText(patient.note),
                            ),
                            Container(
                              alignment: Alignment.center,
```

```
                                    height: tableItemHeight,
                                    width: tableItemWidth,
                                    child: SelectableText(
                                        Tools.birthDateToAge(patient.birthDate)
                                            .toString()),
                                  ),
                                  Container(
                                    alignment: Alignment.center,
                                    height: tableItemHeight,
                                    width: tableItemWidth,
                                    child: SelectableText(patient
                                            .recommendationsCompleted
                                            .toString() +
                                        "/" +
                                        patient.recommendationsCount.toString()),
                                  ),
                                  Container(
                                    alignment: Alignment.center,
                                    height: tableItemHeight,
                                    width: tableItemWidth,
                                    child: SelectableText(
                                        (patient.recentActivityDate == null ||
                                                patient.recentActivityDate == "")
                                            ? "Never"
                                            : Tools.durationAgoString(DateTime.now()
                                                .difference(DateTime.parse(
                                                    patient.recentActivityDate)))),
                                  ),
                                ],
                              ),
                              Container(
                                height: 1,
                                width: double.maxFinite,
                                color: Theme.of(context).dividerColor,
                              ),
                            ],
                          ),
                        ))
                    .toList()),
          ),
        ),
        // ),
      );
    }

    return Scaffold(
      key: scaffoldKey,
      appBar: AppBar(
        title: Text(Constants.applicationName),
        actions: [
          // LOGOUT ICON
          IconButton(
            icon: Icon(Icons.logout),
            onPressed: () => Tools.promptUserLogout(context),
          )
        ],
      ),
      body: Container(
        alignment: Alignment.topCenter,
        child: SingleChildScrollView(
          scrollDirection: Axis.vertical,
          child: SingleChildScrollView(
            scrollDirection: Axis.horizontal,
            child: Container(
              alignment: Alignment.topCenter,
              padding:
                  EdgeInsets.only(left: 130, right: 130, top: 30, bottom: 30),
```

```
                  child: Container(
                    width: Constants.pageMaxWidth,
                    child: Container(
                      child: Column(
                        children: [
                          Container(
                            padding: EdgeInsets.all(16),
                            child: Row(
                              mainAxisAlignment: MainAxisAlignment.spaceBetween,
                              children: [
                                SelectableText(
                                  'Users',
                                  style: Theme.of(context).textTheme.headline4,
                                ),
                                FlatButton(
                                  padding: EdgeInsets.only(
                                      left: 54, right: 54, bottom: 20, top: 20),
                                  color: Theme.of(context).primaryColor,
                                  textColor: Colors.white,
                                  onPressed: _showAddNewPatientModal,
                                  child: Text(
                                    'Add user',
                                    style: Theme.of(context)
                                        .textTheme
                                        .button
                                        .copyWith(fontSize: 16),
                                  ),
                                ),
                              ],
                            ),
                          ),
                          patients.length == 0
                              ? Container(
                                  alignment: Alignment.center,
                                  padding: EdgeInsets.only(top: 250),
                                  child:
                                      _loading // if loading in patients from backend
                                          ? CircularProgressIndicator()
                                          : SelectableText(
                                              'You have no users yet',
                                              style: Theme.of(context)
                                                  .textTheme
                                                  .headline4,
                                            ),
                                )
                              : Card(
                                  child: Column(children: [
                                    tableHeader(),
                                    tableRows(),
                                  ]),
                                ),
                        ],
                      ),
                    ),
                  ),
                ),
              ),
            ),
          ),
        // ),
      );
    }
}
```

C Online Questionnaire about Exercise Trackers

Experiences with workout tracking apps

Page 1

Informed Consent

Background and purpose of the study:

This survey serves as a part of a master thesis project in regards to tendencies with workout and exercise applications. The goal of this experiment is to research personal preferences around different parts of existing workout tracking apps.

The study complies with Norwegian data protection laws and participation in the study is anonymous. Your data will solely be used for research purposes in the specific master project, and after the master thesis study is complete, all data collected will be deleted.

My supervisor is Mariusz Nowostawski and you can find his contact information on https://www.ntnu.edu/employees/mariusz.nowostawski if you have any specific questions to him regarding the study.

Voluntary participation:
Participation in this study questionnaire is fully voluntary. You may choose to withdraw at any point during the study or after the study is completed.

Withdrawing from the study does not require any explanation or reasoning. If you choose to withdraw, all of your data from your participation will be deleted.

If you have any questions or wish to withdraw, you can contact me by e-mail: nikolaaa@stud.ntnu.no

Do you want to participate in this study? *

- ◯ Yes
- ◯ No

Page break

Page 2

Demographics Questionnaire

Age

This element is only shown when the option "Yes" is selected in the question "Do you want to participate in this study?"

- ◯ 18-24
- ◯ 25-30
- ◯ 31-40
- ◯ 41-50

- 51-60
- 61-70
- 71+
- Prefer not to say

Sex

This element is only shown when the option "Yes" is selected in the question "Do you want to participate in this study?"

- Female
- Male
- Prefer not to say

Page break

Page 3

Exercise importance and behaviour

This element is only shown when the option "Yes" is selected in the question "Do you want to participate in this study?"

How important is exercise to you?

	1 (not important)	2	3	4	5	6 (very important)
How important is exercise to you?	○	○	○	○	○	○

How often do you exercise?

This element is only shown when the option "Yes" is selected in the question "Do you want to participate in this study?"

Exercise could be going on a hike, running, bicycling, swimming, skiing, sport activities, etc.

- About every day
- 4-5 times a week
- 2-3 times a week
- Once a week
- Less than once a week
- Never

How hard do you exercise? (on average)

This element is only shown when the option "About every day", "4-5 times a week", "2-3 times a week", "Once a week" or "Less than once a week" is selected in the question "How often do you exercise?"

- ◯ Calmly without becoming sweaty and losing breath
- ◯ So heavy that I become sweaty and out of breath
- ◯ I go all in, such that I have almost no energy left

How long do you exercise each time? (on average)

This element is only shown when the option "About every day", "4-5 times a week", "2-3 times a week", "Once a week" or "Less than once a week" is selected in the question "How often do you exercise?"

- ◯ Less than 15 minutes
- ◯ 15-29 minutes
- ◯ 30-60 minutes
- ◯ Longer than 1 hour

How many hours in a week do you exercise?

This element is only shown when the option "Yes" is selected in the question "Do you want to participate in this study?"

Select …

Page break

Page 4

Exercise preferences

Why do you exercise?

This element is only shown when the option "Yes" is selected in the question "Do you want to participate in this study?"

Tick off each box you agree with.

- ☐ To lose weight
- ☐ To stay in shape
- ☐ Because I want to build muscles
- ☐ Because exercising is fun

- [ ] Because it makes me feel good afterwards
- [ ] Because I compete in sports
- [ ] Because it is a social activity
- [ ] Other

For what other reason do you exercise?

This element is only shown when the option "Other" is selected in the question "Why do you exercise?"

This element is only shown when the option "Yes" is selected in the question "Do you want to participate in this study?"

What type of exercises do you prefer?

	1 (not preferable)	2	3	4	5	6 (very preferable)
Running	○	○	○	○	○	○
Cycling	○	○	○	○	○	○
Swimming	○	○	○	○	○	○
Lifting weights	○	○	○	○	○	○
Climbing	○	○	○	○	○	○
Skiing	○	○	○	○	○	○
Hiking	○	○	○	○	○	○
Sport games (football, badminton, volleyball, tennis, etc)	○	○	○	○	○	○
Gamified workouts (Kinect, Wii, Mobile Apps, VR, etc)	○	○	○	○	○	○

What type of exercises do you prefer?

This element is only shown when the option "Yes" is selected in the question "Do you want to participate in this study?"

If there are any exercises you prefer that were not mentioned in the previous question, please list them here.

Page break

Page 5

Social media and games

How much time do you spend on social media daily? (on average)

This element is only shown when the option "Yes" is selected in the question "Do you want to participate in this study?"

Social media includes platforms such as Facebook, Twitter, Instagram, Snapchat, Web-forums, Reddit, etc.

Select …

Do you play videogames?

This element is only shown when the option "Yes" is selected in the question "Do you want to participate in this study?"

Tick off each platform you play games on.

- ☐ Computer
- ☐ Video game console
- ☐ Virtual Reality
- ☐ Mobile
- ☐ I do not play any videogames

How many hours in a week do you spend playing videogames? (on average)

This element is only shown when at least one of the options "Computer", "Video game console", "Virtual Reality" or "Mobile" are selected in the question "Do you play videogames?"

Select …

Page break

Page 6

Experience with tracking software and technology

Are you using, or have you previously used workout tracking applications?

This element is only shown when the option "Yes" is selected in the question "Do you want to participate in this study?"

For example mobile applications for tracking workouts, apps for smartwatches that track exercises, general health monitoring apps, or similar technologies.

- Currently using
- Have used in the past
- Never used

When using workout tracking apps, how often did you use it?

This element is only shown when the option "Currently using" or "Have used in the past" is selected in the question "Are you using, or have you previously used workout tracking applications?"

- Daily
- Weekly
- Monthly
- Less than monthly

Over how long time period have you used an exercise tracking application?

This element is only shown when the option "Currently using" or "Have used in the past" is selected in the question "Are you using, or have you previously used workout tracking applications?"

- Less than 2 weeks
- 2-4 weeks
- 1-3 months
- 3-6 months
- 6-12 months
- More than a year

Page break

Page 7

Which of the following workout tracking applications do you have experience with?

This element is only shown when the option "Currently using" or "Have used in the past" is selected in the question "Are you using, or have you previously used workout tracking applications?"

- [ ] Strava
- [ ] Garmin
- [ ] Polar
- [ ] Fitbit
- [ ] Google Fit
- [ ] Apple Fitness+
- [ ] adidas Running App
- [ ] MapMyRun
- [ ] Lifesum
- [ ] Zepp
- [ ] Gaming fitness app (Wii, Kinect, VR, etc)
- [ ] Other

What other workout-tracking applications do you have experience with?

This element is only shown when the option "Other" is selected in the question "Which of the following workout tracking applications do you have experience with?"

Which exercise tracking app have you used the most?

This element is only shown when the option "Currently using" or "Have used in the past" is selected in the question "Are you using, or have you previously used workout tracking applications?"

What motivated you to use this particular tracker compared to other trackers?

This element is only shown when the option "Currently using" or "Have used in the past" is selected in the question "Are you using, or have you previously used workout tracking applications?"

Page break

Page 8

For what reasons have you utilized exercise trackers?

This element is only shown when the option "Currently using" or "Have used in the past" is selected in the question "Are you using, or have you previously used workout tracking applications?"

Tick off each box you agree with.

- [ ] It motivates me to exercise
- [ ] I like to analyze my workouts
- [ ] I like to see my personal exercise progress
- [ ] I like to see my personal statistics
- [ ] I like the social aspect (sharing and comparing with friends)
- [ ] Because I like to compete against others
- [ ] Because I like to compete against myself
- [ ] I like to see how many calories I have burned
- [ ] Other

For what other reasons have you utilized exercise trackers?

This element is only shown when the option "Other" is selected in the question "For what reasons have you utilized exercise trackers?"

Which of these health monitoring features have you utilized in your tracker?

This element is only shown when the option "Currently using" or "Have used in the past" is selected in the question "Are you using, or have you previously used workout tracking applications?"

Tick off each feature you have used or are actively using.

- [ ] Sleep tracking
- [ ] Calorie tracking (of food)
- [ ] Eating habits tracking
- [ ] Workout recovery tools
- [ ] Heart-rate monitoring

Page break

Page 9

What fitness tracking technology have you utilized?

This element is only shown when the option "Currently using" or "Have used in the past" is selected in the question "Are you using, or have you previously used workout tracking applications?"

Tick off each box that you have used or actively are using for tracking.

- [ ] Fitness tracking watch
- [ ] Smart watch
- [ ] Mobile app
- [ ] Wrist band
- [ ] Game technology (VR, Kinect, Wii, etc)
- [ ] Exercise machine (Treadmill, rowing machine, exercise bike, etc)
- [ ] Chest strap
- [ ] Heart-rate monitoring
- [ ] Other

What other fitness tracking technology have you utilized?

This element is only shown when the option "Other" is selected in the question "What fitness tracking technology have you utilized?"

Page break

Page 10

What made you start using the exercise tracker(s) you used?

This element is only shown when the option "Currently using" or "Have used in the past" is selected in the question "Are you using, or have you previously used workout tracking applications?"

Tick off each box you agree with.

- ☐ A friend recommended it to me
- ☐ I was bored
- ☐ I wanted to exercise more
- ☐ I wanted to improve the effectiveness of my workouts
- ☐ I wanted to use it to compete with friends
- ☐ I wanted to make my workouts feel more like a game
- ☐ Other

What other reason made you start using the exercise tracker you used?

This element is only shown when the option "Other" is selected in the question "What made you start using the exercise tracker(s) you used?"

Page break

Page 11

This element is only shown when the option "Yes" is selected in the question "Do you want to participate in this study?"

Importance of features in exercise tracking apps

Below are listed several features of exercise tracking apps, rate how important different features are to you if you were to utilize an exercise tracking app.

	1 (not important)	2	3	4	5	6 (very important)
Notifications reminding me to exercise	○	○	○	○	○	○
Graphs displaying my performances in exercises	○	○	○	○	○	○
The ability to set and track personal goals	○	○	○	○	○	○
Being able to compare my perform-ance with others/friends	○	○	○	○	○	○

Being able to see how many calories I have burned in a workout	○	○	○	○	○	○
Receiving workout tips and explanations on how to do specific workouts	○	○	○	○	○	○
Personal workout statistics	○	○	○	○	○	○

This element is only shown when the option "Currently using" or "Have used in the past" is selected in the question "Are you using, or have you previously used workout tracking applications?"

Have you ever been concerned about how your personal data is managed in the application?

Data such as personal workout statistics, personal health, weight, height, etc.

	1 (to a low degree)	2	3	4	5	6 (to a high degree)
Have you ever been concerned about how your personal data is managed in the application?	○	○	○	○	○	○

Is there any type of data you have been more concerned about than others?

This element is only shown when the option "Currently using" or "Have used in the past" is selected in the question "Are you using, or have you previously used workout tracking applications?"

If so, would you like to specify what type of data?

Page break

Page 12

This element is only shown when the option "Yes" is selected in the question "Do you want to participate in this study?"

I think workout tracking software can motivate me to exercise

	1 (to a low degree)	2	3	4	5	6 (to a high degree)
I think workout tracking software can motivate me to exercise	○	○	○	○	○	○

Do you have any NEGATIVE experiences with workout tracking software? Explain.

This element is only shown when the option "Currently using" or "Have used in the past" is selected in the question "Are you using, or have you previously used workout tracking applications?"

Do you have any POSITIVE experiences with workout tracking software? Explain.

This element is only shown when the option "Currently using" or "Have used in the past" is selected in the question "Are you using, or have you previously used workout tracking applications?"

Page break

Page 13

This element is only shown when the option "Currently using" or "Have used in the past" is selected in the question "Are you using, or have you previously used workout tracking applications?"

Exercise tracking app experiences

Below are several statements in regards to exercise tracking apps. Try to the best of your ability to indicate whether they are true or not for your experience with these apps.

	1 (not true at all)	2	3	4	5	6 (very true)
I have previously felt fed up with an exercise tracking app	○	○	○	○	○	○
I have felt annoyed by constant reminders from my exercise app	○	○	○	○	○	○
I think using an exercise tracker improves my workout schedule	○	○	○	○	○	○
I have been concerned about the privacy of my data in an exercise tracker app	○	○	○	○	○	○
I think I am more motivated to exercise by myself than from an app	○	○	○	○	○	○
I have pushed my limits greater in workouts because of an exercise app	○	○	○	○	○	○

I feel good if I see that I perform better than others in the app	○	○	○	○	○	○
Looking at my own previous performances makes me want to beat my own previous records	○	○	○	○	○	○
The app has made me go on a workout I would not have done without it	○	○	○	○	○	○
I feel bad if I see that I perform worse than others in the app	○	○	○	○	○	○

When you used a workout tracking app (now or previously), how many times did you go on a workout because the application influenced you to do so?

This element is only shown when the option "Currently using" or "Have used in the past" is selected in the question "Are you using, or have you previously used workout tracking applications?"

- Almost every day
- A few times a week
- Once a week
- Less than once a week
- Never

This element is only shown when the option "Currently using" or "Have used in the past" is selected in the question "Are you using, or have you previously used workout tracking applications?"

When an app notifies me that a friend of mine has performed better than me, I am motivated to go out and perform better than them in the exercise

	1 (not true at all)	2	3	4	5	6 (very true)
When an app notifies me that a friend of mine has performed better than me, I am motivated to go out and perform better than them in the exercise	○	○	○	○	○	○

Page break

Page 14

Are you still actively using workout tracking apps?

This element is only shown when the option "Currently using" or "Have used in the past" is selected in the question "Are you using, or have you previously used workout tracking applications?"

- Yes
- No

Why did you quit using workout tracker apps?

This element is only shown when the option "No" is selected in the question "Are you still actively using workout tracking apps?"

Tick off each box you agree with.

- The application did not allure to me anymore
- I was tired of working out
- I was annoyed by having constant reminders telling me to work out
- None of my friends were using the application anymore
- I achieved all the goals I wanted to with the application
- The application was boring to use
- I was afraid of how my personal data was handled
- Other reason

For what other reason did you quit using workout tracker apps?

This element is only shown when the option "Other reason" is selected in the question "Why did you quit using workout tracker apps?"

This element is only shown when the option "No" is selected in the question "Are you still actively using workout tracking apps?"

Since you stopped using your workout tracker, how much do you now exercise?

	1 (a lot less than when I used the app)	2	3	4	5	6 (much more than when I used the app)
Since you stopped using your workout tracker, how much do you now exercise?	○	○	○	○	○	○

Is there any specific reason to why you have not used any workout tracking apps before?

This element is only shown when the option "Never used" is selected in the question "Are you using, or have you previously used workout tracking applications?"

If so, explain.

Page break

Page 15

This element is only shown when the option "Yes" is selected in the question "Do you want to participate in this study?"

Study interest

This element is only shown when the option "Yes" is selected in the question "Do you want to participate in this study?"

As previously explained, this questionnaire is part of the author's master thesis work.

He is also seeking participants for a longer study-experiment with some software he has been developing. The study will last over a 2 week period. There will not be activities every day, but a total of 4-6 activities to be done over the course of these two weeks.

In return, you might get some new insights into personal habits and your contributions will help in the present area of research.

Additionally, 4 gift cards (500 NOK each) will be given away in a random raffle to participants of this longer study experiment. Participating will grant you a chance to win one gift card.

Participation will be totally optional and your data will be fully anonymous. If you decide that you are interested in receiving more information about the said study, you will be contacted by me for more information and can on a later date make a decision if you decide to participate or not.

Do you wish to be contacted for more information on the longer study experiment?

This element is only shown when the option "Yes" is selected in the question "Do you want to participate in this study?"

- ○ Yes
- ○ No

Do you have personal access to VR equipment?

This element is only shown when the option "Yes" is selected in the question "Do you wish to be contacted for more information on the longer study experiment?"

- ◯ Yes
- ◯ No

What type of VR equipment do you have access to?

This element is only shown when the option "Yes" is selected in the question "Do you have personal access to VR equipment?"

- ☐ HTC Vive
- ☐ Oculus
- ☐ Playstation VR
- ☐ Smartphone VR
- ☐ Other

What other type of VR equipment do you have access to?

This element is only shown when the option "Other" is selected in the question "What type of VR equipment do you have access to?"

Are you located in the Gjøvik area?

This element is only shown when the option "No" is selected in the question "Do you have personal access to VR equipment?"

If you do not have access to personal VR equipment, it might be possible to conduct experiments at the NTNU Campus in Gjøvik.

- ◯ Yes
- ◯ No

Contact information

This element is only shown when the option "Yes" is selected in the question "Do you wish to be contacted for more information on the longer study experiment?"

If you wish to be contacted to potentially participate in a longer study please contact me on my e-mail: nikolaaa@stud.ntnu.no or describe a way for me to contact you.

Do not use personally identifiable information.

Page break

Survey Submission

If you have decided you do not wish to participate in this study, you can now choose to close your browser and none of your questionnaire answers will be used.

If you still wish to participate, you can submit your survey by pressing the ***send*** button below.

Thank you for your time.

Nikolai Åkerholt

Evaluating a remote data collection platform: A practical user study on exercise trackers

NTNU
Kunnskap for en bedre verden