UNIVERSITY OF OULU
FACULTY OF INFORMATION TECHNOLOGY AND ELECTRICAL ENGINEERING

Jere Koivisto
Aki Korpinen
Lasse Rapo

INDUCING EMPATHY TOWARDS UPPER LIMB IMPAIRMENTS USING A PHYSICAL DEVICE AND VIRTUAL REALITY

Bachelor's Thesis
Degree Programme in Computer Science and Engineering
June 2023

Koivisto J., Korpinen A., Rapo L. (2023) Inducing Empathy Towards Upper Limb Impairments Using A Physical Device And Virtual Reality. University of Oulu, Degree Programme in Computer Science and Engineering, 30 p.

ABSTRACT

Empathy is the ability to understand concepts deeply and intimately from the perspective of another person. Having this empathetic understanding of different medical conditions will help make more informed decisions when designing for a particular condition and increase the motivation for providing higher quality results. However, it can be quite challenging for people to easily gain this kind of empathetic knowledge without fully comprehending the extent to which a particular impairment affects someone's day-to-day life. One of the most popular and effective methods of inducing empathy towards impairments is the use of empathy simulations. The basic concept of empathy simulations is to realistically simulate the limitations posed by an impairment so that the participant can gain a first-hand experience of what it is like to live with the impairment. Traditionally, these simulations were created using various physical means, but lately the use of virtual reality devices in these simulations has become more common. Virtual reality is essentially technology that allows the user to embody another persons perspective, which makes it exceptionally suitable for empathy simulations. The aim of our study was to investigate the generation of empathy towards upper extremity motor impairments using a mixture of physical and virtual means. For the purposes of this study, we built an arm mobility restricting harness to mimic an upper extremity motor impairment and a virtual reality environment of a home kitchen where the simulations took place. Two groups of volunteer participants experienced the simulation by performing simple tasks in the virtual reality environment while being limited by the mobility restricting harness. The difference between the groups was in having to recite different backstories for their simulated characters. Backstory for group 1 was in first-person, and group 2 for group in third-person. The stories were thought to target affective and cognitive empathy differently. The participants' level of empathy was measured once before the simulation and once after the simulation using a collection of standardized questionnaires. The study showed significant increase in the level of emotional contagion over all participants ($p < .044^*$) suggesting that the simulation increased the participants' level of empathy in that category. No significant difference was measured between backstories, however, the results suggest the first-person story to assist cognitive empathy. The study also showed that the group with the backstory in first-person had better scores in all categories of embodiment suggesting that the first-person backstory enabled participants to better relate to their virtual character. Despite some promising results, further studies are needed to investigate empathy generation using a mixed physical and virtual empathy simulations.

Keywords: BSc thesis, empathy, VR, virtual reality, impairment simulation, backstory

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND SYMBOLS

CSE	Computer Science and Engineering
VR	virtual reality
HMD	head mounted display
G	group
H	hypothesis
UEQ-S	Short User Experience Questionnaire
NASA-TLX	NASA Task Load Index
SSQ	Simulator Sickness Questionnaire
QCAE	Questionnaire of Cognitive and Affective Empathy
SUS	System Usability Scale

1. INTRODUCTION

When it comes to designing for disabilities, the key for successful design is empathy. It is the ability to have a deep, intimate understanding of concepts from the perspective of another person. This sort of empathetic understanding of medical conditions helps to make informed decisions when designing around a condition and strengthens the motivation for producing high quality designs. More empathetic medical personnel can design more effective care and rehabilitation, architects can create more accommodating and accessible spaces, and software engineers can create applications that can be enjoyed by everyone regardless of their limitations. However, gaining such empathy can be difficult for people who don't fully comprehend the extents to which an impairment influences day-to-day life.

Numerous kinds of educational material have been devised to spread awareness about common impairments. One of the most popular and effective methods of generating this kind of empathy is via specially crafted impairment empathy simulations. In these empathy simulations, participants are subjected to conditions that try to realistically mimic the limitations of real impairments using a variety of physical and digital means. The goal of these simulations is to generate an intimate understanding about the various effects and challenges the impairment has on day-to-day life by making the participant experience the effects of the impairment firsthand.

The recent advancements in virtual reality technology have made virtual reality applications increasingly more immersive experiences especially in the realm of experimental virtual reality video games. Its applications are not, however, limited to just recreational video games. In fact, the technology is already considered accurate enough for real-life applications that it is used for various training simulations like for training airline pilots. As such there is major potential for using virtual reality technology to enhance the immersion and relatability of empathy simulations.

Virtual reality is essentially technology, that allows the user to embody another person's perspective. This makes virtual reality technology particularly suitable for empathy simulations as the aim is to help the user understand the effects of the impairment from the perspective of having the impairment. Virtual reality technology also allows for a cost effective and flexible way of creating realistic environments for various studies.

This study focuses on the application of virtual reality technology in generating empathy for a family of post-stroke upper extremity motor impairments. To achieve this volunteer participants will go through a relatable virtual scenario of a home kitchen, while wearing an arm mobility restricting harness. The scenario itself has nothing extraordinary about it, but the restriction of arm mobility will provide a realistic challenge for the participants. The participants' level of empathy towards the condition will be measured once before the scenario and once after the scenario.

2. RELATED WORKS

Empathy and virtual reality are not strangers to each other, some even consider VR to be the ultimate tool for empathy [1]. It has been observed that doctor-patient empathy has some measurable effect of success of treatments, while some find empirical evidence lacking[2]. Success of treatments being of great interest in the medical field, ways to induce empathy have been researched and examined and virtual reality has been found to be a potential means to do that [3, 4].

2.1. Empathy

It is important to understand what is meant by empathy, as it can be multidimensional. Generally it is understood to include the ability to understand the situation of another, their perspective, and feelings, as well as communication of one's understanding of them and verify the accuracy [2]. According Dereksen at al., empathy can be understood to have three levels: affective, cognitive and behavioural levels.

When it comes to studies on inducing empathy towards disabilities, studies have managed to induce understanding and empathy towards those suffering from the conditions [3] [4]. However, the studies to enhance empathy can instead increase judgement, as Silverman et al. found that as subjects judge the disabilities to their own, the subjects can view limitations of disabled people as their inherent short comings, rather than inadequacies of the environment [5].

2.2. Upper Limb Impairments

Upper limb impairments can cause significant challenges to ones day-to-day life ranging from wearing clothes to picking up a spoon. Just like the challenges, the conditions are also numerous. One might have limited shoulder movement resulting from an injury while another might have lost dexterity and precision of their hand as a result of a stroke. Arm impairments are quite common side effects of a stroke with issues related to spasticity being present in up to 60 percent of cases [6]. As a result of their frequency, medical field has developed tests to measure the impairments, some of which use day-to-day tasks to evaluate function, while others having subjects perform specific movements and evaluating their performance [7, 8, 9].

2.3. Empathy in Treatments

As mentioned earlier, empathetic relation and interaction between patient and doctor has been of interest in medicine, and research shows mixed results. The general consensus around the subject acknowledges its importance in ensuring patients adhering to instructions and satisfaction on care [2]. In their review, Dersen et al. found that there is some empirical evidence of the benefits to diagnostics, clinical outcomes, even in common diseases like cold, lower anxiety, and patient enabling. They note the need for examination of cost-effectiveness of empathy as that is of interest to policy makers and insurers.

In relation to upper limb impairments, empathy has been shown to affect positively the outcome of botulinum toxin treatment for post-stroke upper limb spasticity [10]. Picelli et al. found significant correlation between empathy and goal attainment scaling, in evaluation of the treatment, but such correlation was missing with other evaluation criteria. They also determine that goal attainment scaling is not to be seen as a clinical measure for spasticity assessment, and should rather be seen as a gauge for symptoms, behaviors, feelings, skills or achievements that a specific treatment is trying to achieve and should be used as a guide for it. Due to the limitations of their study and results indicating that empathy may affect the treatment, they conclude future larger studies, that address the short comings, should be conducted to validate findings.

2.4. Virtual Reality

A simple definition for VR is an immersive virtual 3D environment where the user can operate and interact within [11]. Due to it inherently being immersive, it is different from many other ways to consume and utilize digital content. Its inherently immersive nature has been of great interest for researchers and businesses. Meta, formerly known as Facebook, has invested billions in VR believing there to still be unexplored potential in the technology [12]. As its capabilities are unexplored, it has piqued the interest of some researchers to find means to utilize it.

2.4.1. Emulating Impairments Using Virtual Reality

Virtual reality has been utilized as a tool to emulate a variety of impairments such as impaired vision [13], autism [14], migraine [15], and being wheelchair bound [16]. Commercial entities, charities and research groups have tackled the task of emulating an impairment of their choice, common purpose being to further the understanding of the condition in question. The implementations use a variety of means on how impairments are emulated, for example: making a curated virtual reality experience using processed video footage, limiting how one can maneuver in the environment, or utilizing filters to alter visual experience. These implementations commonly utilize the virtual reality environment to emulate the impairment, whereas our approach on the subject is to use a device to emulate the impairment in the real world.

2.4.2. Virtual Reality and Empathy

The possibilities to use virtual reality experiences to increase empathy has been of interest and has been explored. Dyer et al. developed a virtual reality experience to be used in education of healthcare professionals, and managed to enhance empathy and understanding towards those suffering from age-related health problems, vision and hearing loss, or Alzheimer's disease [3]. In their study to enhance empathy towards visually impaired, Henry et al. achieved increase of empathy towards visually impaired using virtual reality, with his findings being consistent with those of Dyer et al.[4].

Using VR to enchance empathy is not straight forward, since currently there are limitations on how and what deficiencies can be adequately emulated. Yao et al. in their study on empathetic modeling of vision impairment, found that emulating two of the conditions achieved worse results, as subjects were able to cheat the emulated impairment by focusing their vision on clear parts [17]. As a result of their study they found three key design considerations for VR empathy studies: First, ensuring consistency by trying to eliminate variance between subjects resulting from familiarity with VR. Second, preventing cheating by designing the emulation of an impairment in a manner that prevents cheating. And third, maximizing realism, since the realism of the VR environment plays a key role in user immersion.

VR does not necessarily increase empathy as noted by Häyrynen and Pitkänen [18]. In their study on empathy towards color vision deficiency, they observed decrease in empathy when compared to empathy after experiencing the impairment through a computer screen. They note that this could be due to first experience having a greater impact on empathy regardless of the medium, and consider the possibility that VR as an initial experience could have increased empathy more than when experiencing the impairment through a monitor. They also found that both means of experiencing the deficiency increased the understanding of the condition.

2.4.3. Virtual Reality and Rehabilitation

The capability of VR to be an immersive media has been experimented for the purpose of rehabilitation and treatment of various conditions [19, 20]. Anwer et al. in their review article found that studies have found VR to be an effective tool for improving upper limb motor for post-stroke patients. Other factors such as balance, gait, and quality of life also show improvement. However, they assess the evidence level to be weak to moderate, despite positive showings[21]. Noting VR environments being seen as a promising tool in rehabilitation after a stroke, they speculate it to become a key factor in rehabilitation in the near future.

VR has been found to have potential in rehabilitation of children with upper limb motor impairments. In the study conducted by Phelan et al., VR was not only beneficial in improving functional ability, but also was able to reduce perceived pain and perceived difficulty of exercises, enabled increase of exercise duration, and furthered positive attitudes towards therapy [22]. The findings suggest VR to be very effective in the rehabilitation of children.

VR has also been found to be effective in therapeutic intentions on deficits of autism spectrum disorder patients [19]. Karami et al. considers VR to have several advantages, most notable having the ability to provide safe access to realistic environments that would be otherwise considered dangerous. They also note that only VR-based training had strong positive effect on daily living skills of patients, whereas conventional therapy did not, making VR suitable for rehabilitation aimed towards enabling daily living skills.

3. STUDY DESIGN

3.1. Impairment Emulation

To restrict the movement of the subject's arm, our initial design used an exercise resistance band attached to their wrist and an equipment belt. After exploring various ways to limit movement, we found that the resistance band was an affordable and commercially available solution, that did not put a load on the arm in the resting position and provided adjustable resistance.

For the belts initial design, we decided to use the Authorities Crew Belt equipment belt, which is an commercially available belt designed for security professionals to carry their gear. We found that while the resistance band limits the arm movement while extending the arm, the pull on the belt is nearly unnoticeable. In contrast, with traditional belts the pull was noticeable, which could be uncomfortable and unimmersive. The equipment belt is more comfortable due to having a rigid structure as it has been designed to carry a load while not being uncomfortable for its wearer. The load on the belt being minimal serves our purpose of the restricting equipment being less noticeable to the wearer when in VR and minimizes unintended load on them.

We requested feedback on the prototype of the initial design from professionals familiar with the conditions we were trying to emulate. Based on the feedback, the main issue with the initial design was that while it restricted arm movement, the challenge didn't match with the challenge the ones with the impairments have, namely gravity, which is a vertical force, while the restrictive force of our design was lateral. Connection to the wrist was preferred more than connecting the band elsewhere on the arm.

In order to address the issues highlighted in the feedback, we explored possible changes to our implementation. One consideration was to explore other ways to implement the force or restriction other than an exercise band, however that was not the direction we decided further explore. After considering a few ideas to increase the challenge enough, most ideas, such as added weights, seemed impractical to implement or did not satisfy our want to not increase the load when the arm is in a resting position.

The second consideration was to explore other ways to attach the resistance band to create more vertical force. The endpoint of the connection was decided to be on the wrist. There were two means of an attachment to consider: on the person and on the environment. Attaching the band on the environment, for example on the floor, would give us the vertical force we needed. However, such implementations would introduce challenges on how the subject can safely interact and move in the virtual environment, which was not what we desired as the implementation we sought for should have limited restriction beyond emulating the impairment.

The alternative attachments on the body that we explored were located on the lower limb, on the same side as the restricted hand. After considering and testing possible ways to attach the band on the foot, ankle, shin, knee and thigh, attachment to lower thigh, slightly above the knee was found to be the most suitable for our purposes. Attachments below the knee, while providing more vertical challenge, had a fundamental issue of affecting the walking and stability of the subject, and thus were ruled out of consideration. The attachment on the knee, while otherwise similar to the one on the thigh right above, had a small but noticeable effect when bending the knee. While the effect was minor, it was noticeable enough that it could have had an undesirable effect on the subjects' immersion in VR. Therefore, the attachment on the lower thigh, above the knee, was chosen.

3.1.1. Restriction Equipment

Our implementation of the design uses an elastic belt around the thigh and a weightlifting wristband around the wrist connected with an exercise band (see Figure 1). Both, the belt and wristband, can be adjusted for adequate tightness on the subject and are easy and quick to wear and remove by either the subject or another person. For the exercise band, we chose a commercially available band that we found to provide adequate challenge.

In order to better attach the band between the belt and the wristband, we decided to design and produce our own solution as suitable commercial solutions were not available. We used Autodesk Fusion 360 to create 3D models of connector pieces that were attached to the wristband and the belt, and the exercise band would then be connected between the attachments. The connector of the wrist has a loop for looping the exercise band through it, and the band is then tied either with a knot or a zip tie. The connector on the thigh has a double loop design for tying the band with a self-

Figure 1. Example of a person wearing the device.

tightening knot, making the connection reliable and enabling easy adjustment of the length of the connection. The parts were 3D printed with Raiser3D Pro2 using ABS filament, ABS having more suitable properties than PLA.

When producing the piece connected to the belt on thigh using sparse infill density, it was observed that it introduced some flex to the part. This combined with strong pulling force could lead to outer layer to separate from infill around the corners of the hook leading to structural compromise. The issue was addressed by using denser infill, eliminating sharp corners, and slightly increasing thickness to avoid flex.

3.2. VR Hardware

To create an immersive experience for the participants and obscure the physical setup, we chose to use the HTC Vive VR as our head-mounted display (HMD). The HTC Vive VR system is a high-end virtual reality platform developed by Valve Corporation and HTC in 2016. This HMD provides two accompanying hand-held controllers that allow the user's movement and active engagement with the virtual environment. For our purposes, the VR system was more than sufficient despite being an older model. It's compatible with different computer setups and comfortable to use. The HMD and controllers were disinfected after every participant with an UV light.

3.3. Software

The software for creating virtual scenarios needed to be flexible, fast to use and powerful enough to handle our VR scenarios. We ultimately chose to use Unity since we had earlier personal experience of using it, and it has a plenty of documentation for creating VR scenarios. Unity's support in different VR hardware also makes it a safe and efficient option for our use. Unity uses C# as its primary programming language. It also has a vast collection of free 3D assets and plugins which allowed us to quickly prototype different scenarios.

3.4. VR Scenario

The aim of the study is to have the test subjects go through a relatable scenario in which they can experience the challenges imposed by the mobility impairment. The virtual environment had a major role in creating a realistic and immersive experience for the test subjects. In order to show the test subjects the contrast in challenge brought by the impairment, the virtual environment and the tasks within it had to mimic common situations the type of which the test subject has likely already found themselves in before.

Figure 2. Virtual environment of a home kitchen.

Because of this, we decided to base our virtual scenario in a home kitchen setting, as shown in Figure 2 and Figure 3, with the task of making a simple cheese pizza. The home kitchen is a particularly relatable setting for most test subjects and provides many opportunities for hand- straining tasks in the form of cooking a meal. Pizza was chosen as the dish in question because it provides straightforward recipes that can be easily followed by test subjects. It is important to keep the tasks within the virtual scenario simple in order to ensure that the challenge comes not from the scenario, but the mobility restriction imposed by the harness.

The scenario is built using the Unity game development tools and is designed to run on top of SteamVR using the HTC Vive virtual reality kit. In the scenario, the test subject can grab objects and move around by teleporting using the VR controllers given to them. A set of tasks is created from a simple cheese pizza recipe that the test subject must then follow. The set of tasks contains the following:

- Rolling the pizza dough with a rolling pin.
- Pouring tomato sauce on the flat pizza pie and spreading it with a wooden spoon.
- Grating a block of cheese using a cheese grater and spreading the grated cheese on the pizza.
- Putting the pizza into an oven.
- Taking the now cooked pizza out of the oven.

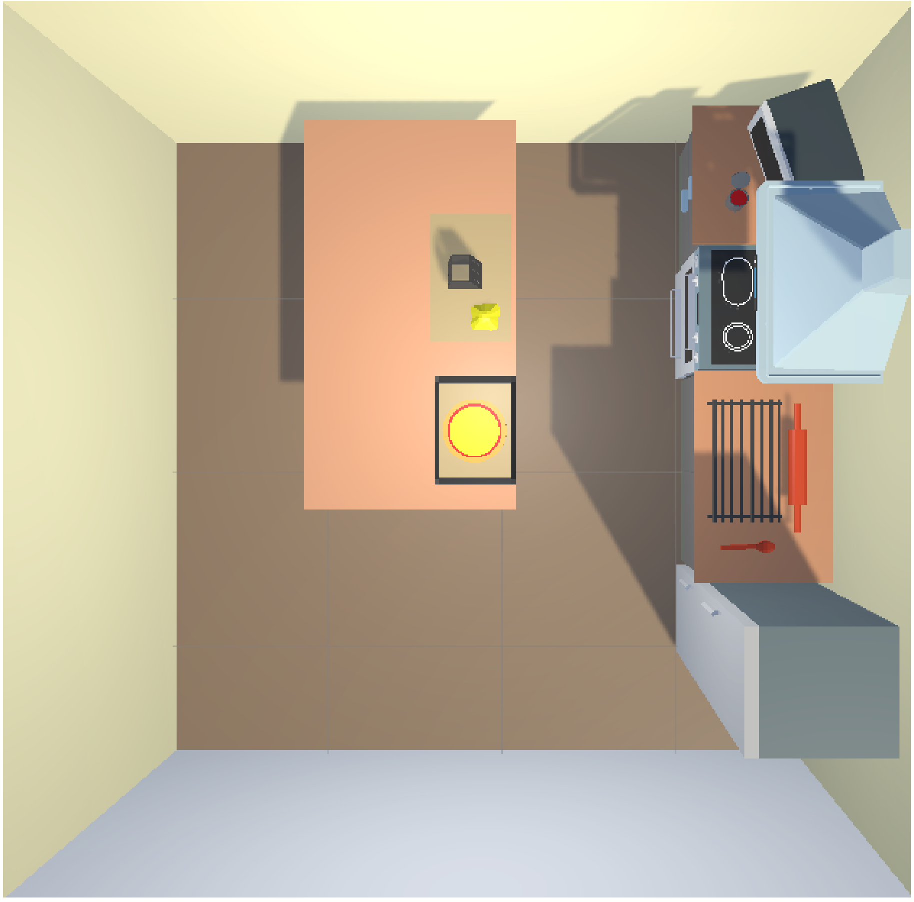

Figure 3. View from above.

The scenario works by establishing an integer state and a list of tasks within a simple switch. Once a task is completed, the state is incremented, and the program switches to updating the next task. In order to complete the tasks, the test subjects must manipulate objects by grabbing them and using said objects to perform actions. All of the above tasks have been designed to be trivial to complete with unrestricted hand mobility. This ensures that the challenge comes only from the imposed mobility restriction.

Most of the tasks involve making some type of motion using various interactable items (see Figure 4), such as grating cheese and rolling dough with a rolling pin. All of these actions work by measuring the total distance moved by a given object within an assigned trigger volume. Once the total moved distance has exceeded a given threshold, an if statement will change the current task into a new one. By modifying the shape and size of the triggers, we can create tasks that require different motions to be performed compared to other tasks. For example, the dough has a wide disc shaped trigger for long flat strokes with the rolling pin, whereas the cheese grater has a small sphere-shaped trigger for small and sharp upwards strokes.

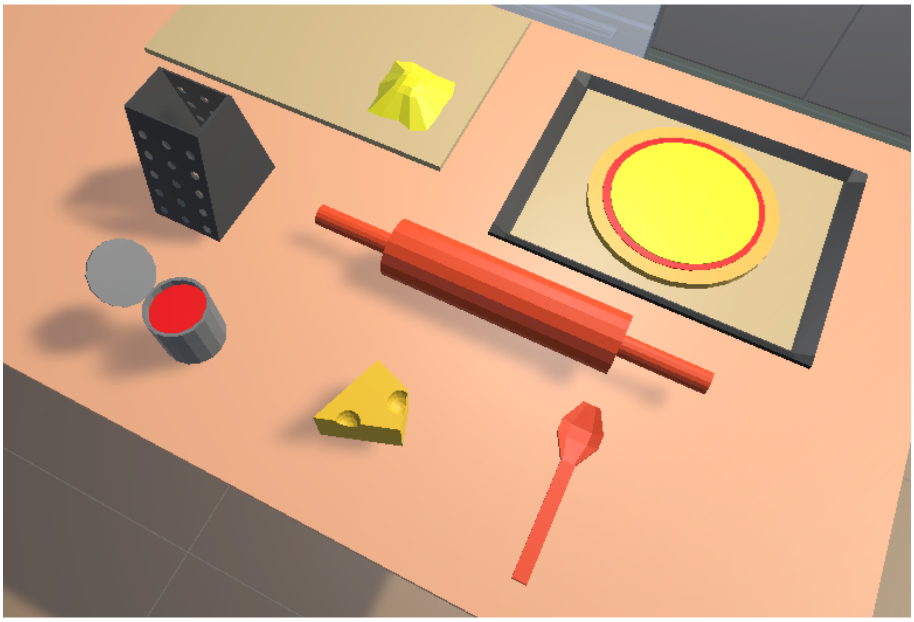

Figure 4. Various items the test subject interacts with.

The virtual environment has been built with a mixture of premade assets from the Unity asset store as well as 3D models and scripts purposefully created for this study. Most of the common VR interactions such as grabbable physics objects and teleport movement as well as compatibility with SteamVR came from Valves' official SteamVR plugin [23] that was downloaded from the Unity asset store. Other premade assets include a clean low poly set of kitchen appliances [24] and cabinets [25]. The rest of the 3D models such as models for interactable objects were created in 3D modeling tool Blender for use in this study.

3.5. Experiment Process

The experiment was conducted in the premises of Oulu University in a room big enough for participants to freely walk around during the simulation (see Figure 5). It is important to keep the experiment process as similar as possible between every participant to eliminate any unwanted factors that could affect the results. Two to three researchers were present in the room during each experiment session. Only one researcher communicates with the participant during the experiment. This approach ensures that the participant doesn't feel uncomfortable or confused during the experiment. It also minimizes any sources of bias which leads to more honest and accurate responses from the participant, improving the quality of the data collected. The participants were split into two different groups (G) by using two different backstories.

- G1 reads backstory 1.
- G2 reads backstory 2.

Figure 5. Study setup.

The experiment process goes as follows (see Figure 7):

- The researchers are introduced to the participant and the participant is seated (see Figure 6).
- The participant reads and signs consent forms.
- The experiment process is briefly explained to the participant.
- The participant answers the first questionnaire.
- The participant reads a short guide of the VR environment. The printed guide shows the tools and their names used in the VR simulation to help non-english speakers better identify them.
- The participant reads aloud a backstory of the role they are about to play in the VR simulation.
- The mobility limiting harness is put on the participant and the participant goes through the VR simulation.
- The participant answers the second questionnaire.

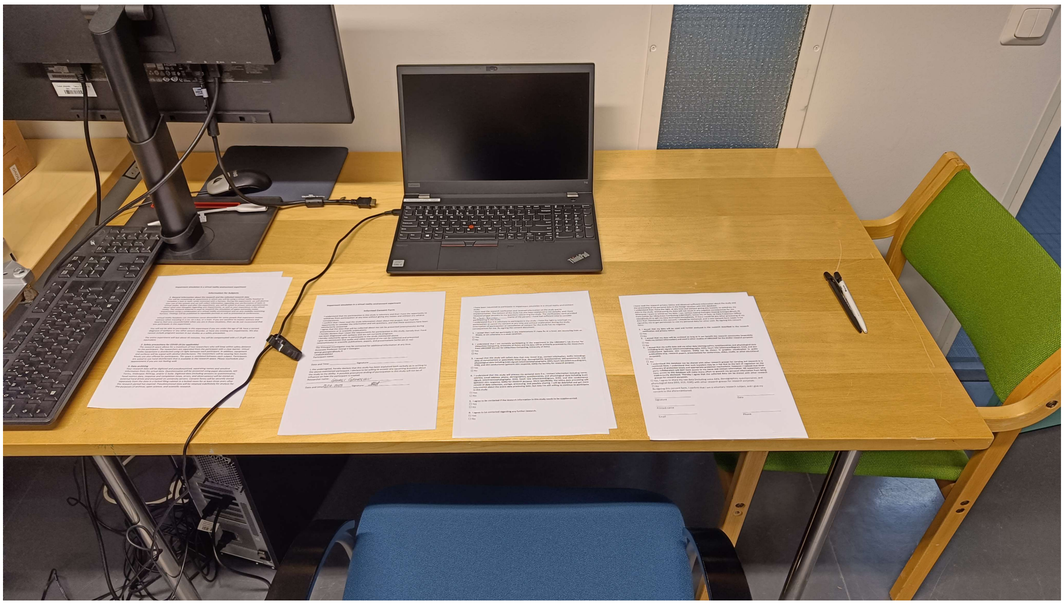

Figure 6. A table and a laptop for answering the questionnaires.

3.5.1. Backstory

Presenting a backstory to an individual has the potential to increase empathy according to the Affective Disposition Theory, making the use of a backstory a valuable strategy for promoting empathy [26]. The primary difference between the two backstories is the focus of the narrative. The first backstory is aimed to induce affective empathy, while the second backstory is aimed to induce cognitive empathy.

The first backstory depicts a person who has suffered a stroke six months prior and is preparing to attend their friend's potluck birthday party by making a pizza to share. It emphasizes the physical and emotional struggles that the person faces. The backstory is written in first-person with the stroke survivor as the narrator, to create a sense of personal connection and emotional connection, which in turn can lead to a greater sense of empathy towards the character and the difficulties they face. A first-person narrative can help to humanize the character and make them feel more relatable, which can also increase empathy [27].

The second backstory is similar to the first backstory in terms of content, but focuses on the cognitive challenges that the person must overcome. It is written in the third-person view, with the narrator describing the experiences and thoughts of the stroke survivor from and external perspective, to create and intellectual understanding of the character's experiences.

3.6. Aim and Hypothesis

To ensure that the questionnaires are aligned with the research goals, we have to identify the key questions that we aim to investigate. The main things we want to examine are:

1. Will the virtual environment combined with the physical setup help to enhance empathy towards individuals with the real condition?
2. Will the virtual environment help to effectively immerse the test subjects in the role of a person with the condition by hiding the physical setup?
3. How different backstories affect gained empathy from the VR experience?

To address these questions, we have formulated the following hypotheses:

H1. The combination of the virtual environment and the physical setup will help the test subjects to gain more empathy towards individuals with the real condition.
H2. The virtual environment will enhance immersion and enables the test subjects to take the role of a person with the real condition.
H3. The groups that are given different backstories will have different changes in empathy.

3.7. Questionnaires

The questionnaires are split in to two different parts. The first questionnaire is filled by the participant before the VR simulation. The Pre Simulation Questionnaire contains the following sections:

- Demography info: participant ID, age and gender. The unique participant ID is given by the researcher and is used to match the two questionnaires.
- Participant's VR experience. Helps us to understand how much the physical restriction is affecting the difficulty of the simulation.
- Simulator Sickness Questionnaire (SSQ). Gives us the baseline of the wellbeing of the participant to compare with the post simulation results. [28]
- Questionnaire of Cognitive and Affective Empathy (QCAE). Compared with the post simulation results to see how the simulation affects the participants empathy. [29]

The second questionnaire filled by the participant after the VR simulation. The Post Simulation Questionnaire contains the following sections:

- Demography info: participant ID.
- Questionnaire of Cognitive and Affective Empathy (QCAE).
- Simulator Sickness Questionnaire.
- NASA Task Load Index (NASA-TLX). Used to measure the perceived workload of the simulation. [30]
- Presence Questionnaire. Measures the subjective sense of "presence" in the virtual environment. [31]

- Embodiment Questionnaire. Measures the subjective sense of embodiment or ownership of the virtual body/hands. [32]
- Short User Experience Questionnaire (UEQ-S). Covers a comprehensive impression of user experience. [33]
- System Usability Scale (SUS). Provides a reliable tool for measuring the usability of the simulation. [34]
- Goodness of Fit. A few questions of our own design to get feedback on how good and easy to use the gestures used in the virtual environment were, and how challenging the harness was.

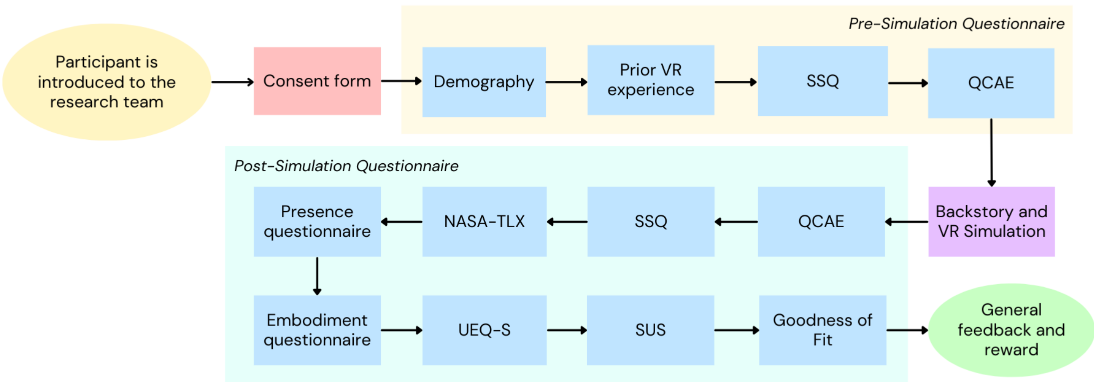

Figure 7. Flowchart of the experiment process.

4. RESULTS

4.1. Participant Demographics

The study had total of 27 participants, 14 of which were randomly placed in group 1 and the other 13 in group 2, with differing backstory. 63% of the participants were men, 37% were women. The majority of the participants were 25 to 34 years old, and many of the participants were students or researchers at the University of Oulu. Nearly all of the participants had prior experience in a VR study.

4.2. Empathy

The questions were answered in 5-point likert scale instead of the suggested 4-point scale. For the analysis, the answers were converted from a 5-point scale to a 4-point scale [35]. This adjustment is for the purposes of analysis only, and it is possible that having a neutral option had an effect on the results.

The pre- and post-questionnaire answers were scored according to the design of Reniers et al.[29]. Affective and cognitive empathy consists of the sum of results in subcategories as follows: emotion contagion, proximal responsivity and peripheral responsivity make up affective empathy, and perspective taking and online simulation make up cognitive empathy. The results were analyzed for each group using a paired t-test, comparing results before and after the VR experience. Between-group analysis was done with a two-sample t-test, considering the changes for each subject as the values for each category.

While most of the results did not reach significant *p* values, change in empathy can be observed and some indications of possible differences between groups can be inferred from results shown in Table 1. Group 1 had slight increase in empathy and group 2 experienced slight decrease, however the results are not statistically significant. Most significant improvement can be observed in emotion contagion, with groups individually reaching significance of $p < .174$ and $p < .128$, and for the subjects as a whole the improvement reaches $p < .044$ which indicates significant result.

	Affective empathy	Cognitive empathy	Emotion contagion	Proximal responsivity	Peripheral responsivity	Perspective taking	Online simulation	Total
Group 1 n=14	0.214 $p < .650$	0.321 $p < .685$	0.268 $p < .174$	0.214 $p < .526$	-0.268 $p < .615$	0.375 $p < .545$	-0.054 $p < .901$	0.536 $p < .574$
Group 2 n=13	0.519 $p < .583$	-0.635 $p < .276$	0.635 $p < .128$	0.053 $p < .901$	-0.173 $p < .673$	-0.231 $p < .559$	-0.404 $p < .415$	-0.115 $p < .926$
Between groups	0.305 $p < .765$	-0.956 $p < .332$	0.367 $p < .392$	-0.157 $p < .780$	0.095 $p < .887$	-0.606 $p < .412$	-0.350 $p < .585$	-0.651 $p < .671$
All n=27	0.361 $p < .472$	-0.139 $p < .776$	0.444 $p < .044$*	0.139 $p < .614$	-0.222 $p < .500$	0.083 $p < .820$	-0.222 $p < .483$	0.222 $p < .769$

Table 1. Average changes in empathy by category for each group and the *p* values.

The main difference between groups, and why the other had overall decrease in empathy while the other did not, is due to difference in change of cognitive empathy, especially perspective taking. While none of the differences were revealed to be of statistical significance by our analysis using a two-sample t-test, it could suggest a possible positive effect of the first-person backstory.

4.3. User Experience

The overall user experience of the participants is measured by the SSQ, the Embodiment questionnaire, the UEQ-S, the SUS, the Goodness of Fit questionnaire and the NASA-TLX questionnaire. The Presence questionnaire also measures the user experience and user involvement in the virtual environment, but is excluded from further analysis due to high variability in the data (see Figure 8).

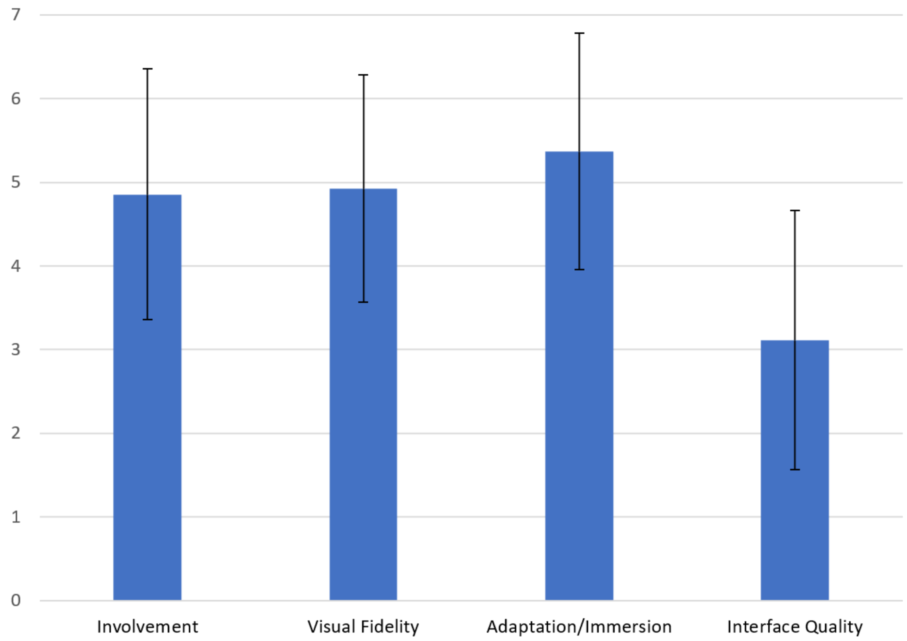

Figure 8. Mean scores and standard deviations for each factor of the Presence questionnaire. Audio fidelity was not measured since the VR simulation did not have any audio.

4.3.1. Embodiment Questionnaire

The test subjects were tasked with answering the questions of the Avatar Embodiment questionnaire during the post experiment questionnaire. The results of the questionnaire (see Figure 9), seem to suggest moderately better embodiment scores in all categories for the group who were given the first-person backstory targeting affective empathy as opposed to third-person backstory targeting cognitive empathy. This result could suggest that the more emotionally resonant story told from the first perspective had a positive impact on how well the test subject was able to relate to the player character in the simulation. The lower results in appearance and response categories compared to other categories, which could suggest the primary issue in embodiment having been the virtual environment, and not the restriction device.

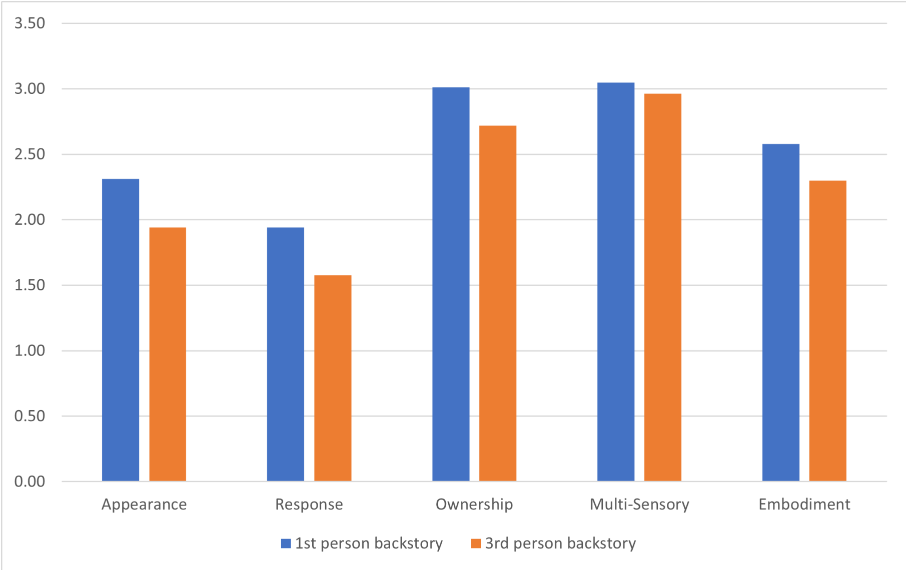

Figure 9. Mean scores for each of the categories from the embodiment questionnaire.

4.3.2. UEQ-S, SUS and Goodness of Fit

The UEQ-S is split into a number of task-related (pragmatic) and non-task-related (hedonic) user experience aspects. The questionnaire has seven items in a 7-point likert scale (see Table 2). The first four items measure pragmatic quality and the last three items measure hedonic quality.

Table 2. The items used in the UEQ-S

Negative	1	2	3	4	5	6	7	Positive
obstructive								supportive
complicated								easy
inefficient								efficient
confusing								clear
boring								exciting
not interesting								interesting
conventional								inventive

The data of the UEQ-S is transformed to a scale of -3 (horribly bad) to +3 (extremely good). The average of those values then represent the quality of an item. The results of the UEQ-S show an overall positive trend (see Figure 10). Values > 0.8 represent a positive evaluation (green), values between 0.8 and -0.8 represent a neutral evaluation (yellow), values < -0.8 represent a negative evaluation (red). The most positive item, with a mean of 2.2, was the clearness of the simulation. This was likely thanks to the clear instructions given to the participants before and during the simulation. The items with the lowest mean of 1.2 were supportiveness and inventiveness.

The original SUS scale has ten items used for global assessments of systems usability. We used eight items of the SUS suitable for measuring the usability of the simulation, making the SUS score have a range of 0 to 80, with a higher score being better. The average score was 60.28, indicating that the participants didn't have any greater problems during the simulation.

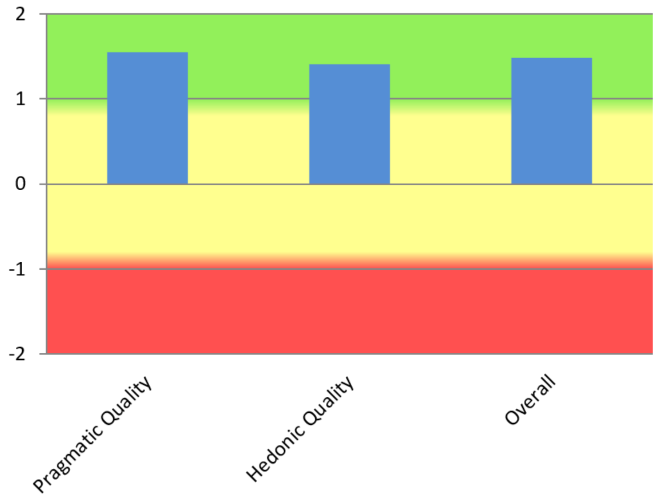

Figure 10. UEQ-S scales: pragmatic quality = 1.546, hedonic quality = 1.407, overall quality = 1.487.

The Goodness of Fit questionnaire had three items that are more general compared to those in the UEQ-S. They are also answered by using a 7-point likert scale. The Goodness of Fit questionnaire had quite similar results to the UEQ-S. Participants felt that the VR environment and harness were well made (see Figure 11). In terms of challenge, the harness was rated somewhat challenging.

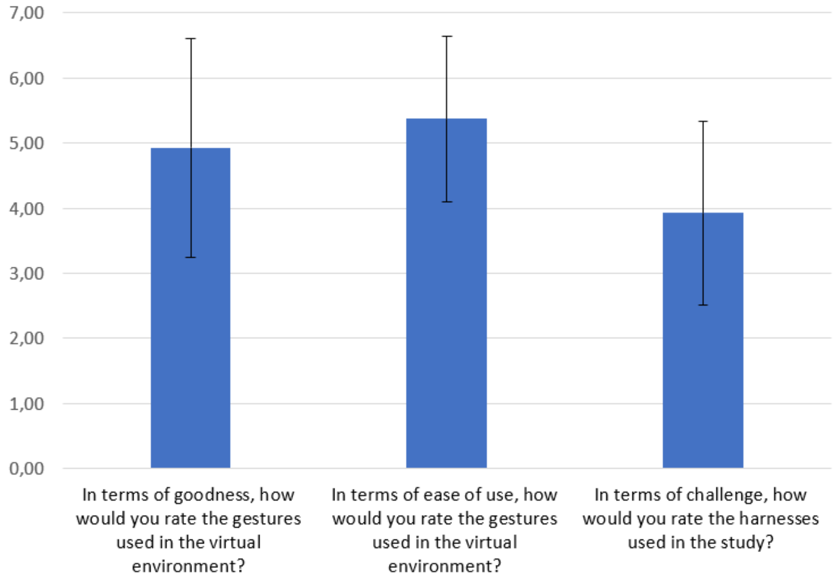

Figure 11. Mean scores and standard deviations for each item of the Goodness of Fit questionnaire.

4.3.3. NASA-TLX

Several types of demand caused by the simulation were measured using the NASA-TLX questionnaire to which the test subjects answered during the post experiment questionnaire. The results of this questionnaire (see Figure 12) seem to suggest that the participants who were given the third-person backstory found the simulation on average more demanding in almost all categories except for temporal demand. The most notable increases were in how much effort the test subjects felt that they had to give in order to complete the tasks and in how frustrated they felt while completing the tasks. Although despite feeling more frustrated with the tasks, the group also felt on average that their performance in completing the tasks was somewhat greater.

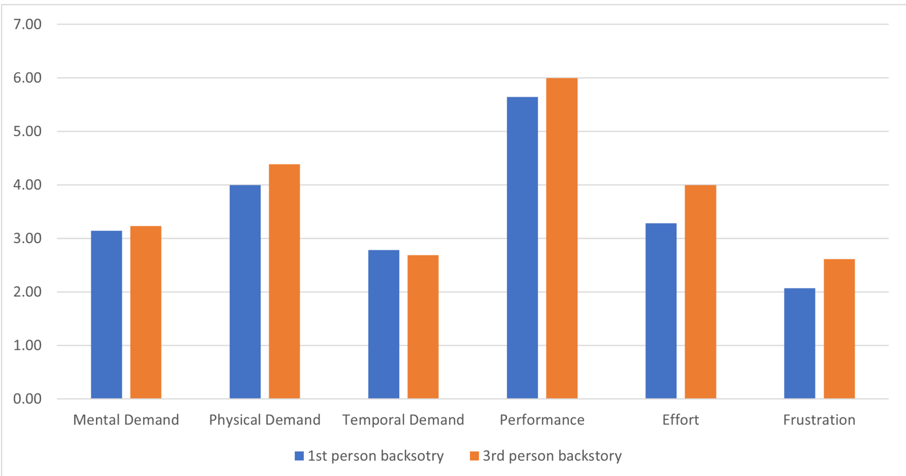

Figure 12. Mean scores for each of the categories from NASA-TLX questionnaire.

4.3.4. Simulator Sickness Questionnaire

To evaluate whether the simulation caused any discomfort or sickness to the participants, we used the SSQ, which evaluates the sickness elicted by virtual reality systems. The results of the SSQ suggest that the simulation did not cause any additional significant discomfort to the participants compared to a typical use of an HMD. The majority of symptoms affecting the participants were related to eye strain and blurred vision, and in most cases these symptoms were experienced by individuals who wear eyeglasses. Six participants felt slight fatigue.

5. DISCUSSION

5.1. Changes in Empathy

Most of the observed change of empathy in our study did not reach significant levels, however, change in emotional contagion over all participants reached significant levels of $p < .044$, indicating potential in our approach to the subject. The differences between groups also did not reach significant levels, however, some potential was indicated for using first-person story over third-person one to assist the increase of empathy, or to combat its reduction.

QCAE was found to be a valuable tool to further look into how empathy was affected in our study. However due to study set up some issues were observed. Use of 5-point scale could have had some effect on results, and based on participant interviews, the proximity of the two questionnaires could be an issue, as some could and tried recall their first answers when answering for the second time. This could be addressed by increasing the time between, for example, instead of having both in one experiment session, the first questionnaire could be at the time of registration to participate.

Due to aforementioned issues and relatively low significance levels, the results should not be considered definitive proof of successes or failure of our approach, and should be seen as indications of where possible benefits of our approach lie for future works.

5.2. Hypotheses

Considering our hypotheses against the results, assuming the change of empathy being a result of experiencing impairment in VR, and that change being towards the condition, the significant increase of emotional contagion suggests that our first hypothesis is confirmed regarding that subcategory of empathy.

Our second hypothesis of VR enabling immersion was inconclusive. Participants were able to embody their virtual avatar while wearing the restriction harness and experienced increase in empathy, however, further examination is needed to confirm whether this was due to the usage of VR or some other factor such as the backstory.

Our last hypothesis was that groups with different backstories would have differing changes in their level of empathy. This hypothesis was ultimately unconfirmed as we found no significant difference in the increase of empathy for the two groups.

5.3. Future Work

We found that in our study there is improvement to be made both in the physical and in the virtual setup. The physical setup should mimic the real-life condition or situation as accurately as possible that we are simulating. In our case an issue that caused some confusion was the fact that the resistance band also restricted the movement of the leg that the band was attached to, especially with smaller participants. This issue was especially prominent when the resistance band was attached tighter. The attachment to the leg also allowed participants to lift their leg to help with the tasks. Such cheating was observed among several participants.

These issues were more or less visible between different participants, so the physical setup should have a good method for individualizing the challenge depending on the characteristics of the participant. Based on the feedback from some participants and professional therapists was also that upper limb impairment caused by a stroke should limit the hands movement a lot more than the resistance band does. The harness did not impose enough challenge to be comparable to the real challenge faced by real stroke survivors.

Some participants gave feedback that the virtual environment looked unrealistic or felt unresponsive. In the future, when a hyper-realistic VR-environment can be used, it could be studied if it allows for an even more empathetic experience. Additionally, more emphasis should be put on correct response to actions during tasks, especially as the virtual environment relied on participants' lived experience in tasks.

6. CONCLUSIONS

The aim of this study was to asses the effectiveness of a combined virtual reality and physical limitation setup in terms of generating empathy towards a target impairment. This was researched by creating a study in which volunteer participants take part in an impairment simulation consisting of a virtual reality environment and an arm mobility restricting harness. The participants' level of empathy was measured using questionnaires before and after the simulation. The study also examined how two different backstories affected gained empathy.

We designed a device to emulate upper limb impairment, using mostly off-the-shelf parts. We also made a virtual environment, which utilized participants living experience to draw comparisons to. The virtual environment performed well for the purpose based on questionnaire feedback.

We observed significant change $p < .044$ in emotion contagion over all participants, while change in other categories did not reach statistically significant levels. Results between groups did not have a significant difference, however observed differences suggest that the first-person backstory helps to induce cognitive empathy. The backstories were observed to have some effect on embodiment, as the group with the first-person story on average scored higher in embodiment, suggesting it assists participants in relating to the virtual body.

Further improvements to the study could be made. For example, further emphasis could be put in creating a simulation that allows for more realistic feedback from interactions. A longer time period could be left between the QCAE questionnaires, to avoid participants trying to recall their previous answers.

7. REFERENCES

[1] Chris Milk: How virtual reality can create the ultimate empathy machine | TED Talk. URL: https://www.ted.com/talks/chris_milk_how_virtual_reality_can_create_the_ultimate_empathy_machine.

[2] Derksen F., Bensing J. & Lagro-Janssen A. (2013) Effectiveness of empathy in general practice: A systematic review. British Journal of General Practice 63.

[3] Dyer E., Swartzlander B.J. & Gugliucci M.R. (2018) Using virtual reality in medical education to teach empathy. Journal of the Medical Library Association 106, pp. 498–500.

[4] Henry T., Qazi S., Kutanzi E., Damji K.F., Whitford R. & Al Hamarneh Y.N. (2021) Enhancing medical professionals' and students' empathy for visually impaired patients using virtual reality. Canadian Journal of Ophthalmology 56, pp. 338–339.

[5] Silverman A.M., Gwinn J.D. & Van Boven L. (2014) Stumbling in Their Shoes. http://dx.doi.org/10.1177/1948550614559650 6, pp. 464–471. URL: https://journals.sagepub.com/doi/abs/10.1177/1948550614559650?journalCode=sppa.

[6] Chohan S.A., Venkatesh P.K. & How C.H. (2019) Long-term complications of stroke and secondary prevention: an overview for primary care physicians. Singapore Medical Journal 60, p. 616. URL: http:/dx.doi.org/10.11622/SMEDJ.2019158.

[7] Upper Extremity Function Test (UEFT) – Strokengine. URL: https://strokengine.ca/en/assessments/upper-extremity-function-test-ueft/.

[8] Arm Motor Ability Test - Physiopedia. URL: https://www.physio-pedia.com/Arm_Motor_Ability_Test.

[9] Wolf Motor Function Test (WMFT) – Strokengine. URL: https://strokengine.ca/en/assessments/wmft/.

[10] Picelli A., Vallies G., Chemello E., Gavras A., Castellazzi P., Meschieri A., Serina A., Gandolfi M., Baricich A., Santamato A., Cisari C. & Smania N. (2017) Influence of physician empathy on the outcome of botulinum toxin treatment for upper limb spasticity in patients with chronic stroke: A cohort study. JOURNAL OF REHABILITATION MEDICINE 49, pp. 410–415. URL: https://fair.unifg.it/handle/11369/357811.

[11] Sheldon R., What is virtual reality? URL: https://www.techtarget.com/whatis/definition/virtual-reality.

[12] Dean G. (2022), Meta has pumped $36 billion into its metaverse and VR businesses since 2019. These 4 charts show the scale of its extreme spending — and huge losses. URL: https://www.businessinsider.com/charts-meta-metaverse-spending-losses-reality-labs-vr-mark-zuckerberg-2022-10?r=US&IR=T.

[13] Virtual Reality Visual Impairment Experience — Angel Eyes NI. URL: https://www.angeleyesni.org/vrviexperience.

[14] Autismity - The Autism Simulator. URL: https://theautismsimulator.com/.

[15] Misztal S., Carbonell G., Zander L. & Schild J. (2020) Simulating Illness: Experiencing Visual Migraine Impairments in Virtual Reality. 2020 IEEE 8th International Conference on Serious Games and Applications for Health, SeGAH 2020 .

[16] Wheelchair Simulator VR. URL: https://www.oculus.com/experiences/rift/1655593951188670/.

[17] Yao T., Yoo S. & Parker C. (2021) Evaluating Virtual Reality as a Tool for Empathic Modelling of Vision Impairment: Insights from a simulated public interactive display experience. ACM International Conference Proceeding Series , pp. 190–197.

[18] Häyrynen E. & Pitkänen S. (2022) Empathy and idea generation for color vision deficiency in virtual reality. Ph.D. thesis.

[19] Karami B., Koushki R., Arabgol F., Rahmani M. & Vahabie A.H. (2021) Effectiveness of Virtual/Augmented Reality–Based Therapeutic Interventions on Individuals With Autism Spectrum Disorder: A Comprehensive Meta-Analysis. Frontiers in Psychiatry 12, p. 887.

[20] Mouawad M.R., Doust C.G., Max M.D. & McNulty P.A. (2011) Wii-based movement therapy to promote improved upper extremity function post-stroke: A pilot study. Journal of Rehabilitation Medicine 43, pp. 527–533.

[21] Anwer S., Waris A., Gilani S.O., Iqbal J., Shaikh N., Pujari A.N. & Niazi I.K. (2022) Rehabilitation of Upper Limb Motor Impairment in Stroke: A Narrative Review on the Prevalence, Risk Factors, and Economic Statistics of Stroke and State of the Art Therapies. Healthcare 10.

[22] Phelan I., Furness P.J., Matsangidou M., Carrion-Plaza A., Dunn H., Dimitri P. & Lindley S.A. (2021) Playing your pain away: designing a virtual reality physical therapy for children with upper limb motor impairment. Virtual Reality URL: https://doi.org/10.1007/s10055-021-00522-5.

[23] Corporation V., SteamVR Plugin for Unity. URL: https://assetstore.unity.com/packages/tools/integration/steamvr-plugin-32647.

[24] Infinite A., Kitchen Appliance - Low Poly. URL: https://assetstore.unity.com/packages/3d/props/electronics/kitchen-appliance-low-poly-180419.

[25] Infinite A., Kitchen Cabinets - Low Poly. URL: https://assetstore.unity.com/packages/3d/props/interior/kitchen-cabinets-low-poly-183890.

[26] Shady M.A. (2020), The relationship between reading, character morality, and narrative empathy. Undergraduate senior honors thesis, Ball State University. URL: http://cardinalscholar.bsu.edu/handle/123456789/202588.

[27] Shi Y. (2021) First-Person Narrative and Story Meaningfulness: Promoting Empathy via Storytelling. Master's thesis, University of Chicago. URL: https://knowledge.uchicago.edu/record/2864.

[28] Kennedy R.S., Lane N.E., Berbaum K.S. & Lilienthal M.G. (1993) Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. The International Journal of Aviation Psychology 3, pp. 203–220.

[29] Reniers R.L., Corcoran R., Drake R., Shryane N.M. & Völlm B.A. (2011) The QCAE: A questionnaire of cognitive and affective empathy. Journal of Personality Assessment 93, pp. 84–95.

[30] Hart S.G. & Staveland L.E. (1988) Development of nasa-tlx (task load index): Results of empirical and theoretical research. In: P.A. Hancock & N. Meshkati (eds.) Human Mental Workload, Advances in Psychology, vol. 52, North-Holland, pp. 139–183. URL: https://www.sciencedirect.com/science/article/pii/S0166411508623869.

[31] Witmer B.G., Jerome C.J. & Singer M.J. (2005) The factor structure of the presence questionnaire. Presence: Teleoperators and Virtual Environments 14, pp. 298–312. URL: https://dl.acm.org/doi/10.1162/105474605323384654.

[32] Peck T. & Gonzalez-Franco M. (2021) Avatar embodiment. a standardized questionnaire. Frontiers in Virtual Reality 1.

[33] Schrepp M., Hinderks A. & Thomaschewski J. (2017) Design and evaluation of a short version of the user experience questionnaire (ueq-s). International Journal of Interactive Multimedia and Artificial Intelligence 4, p. 103.

[34] Brooke J. (1995) SUS: A quick and dirty usability scale. Usability Eval. Ind. 189.

[35] Transforming different Likert scales to a common scale. URL: https://www.ibm.com/support/pages/transforming-different-likert-scales-common-scale.