Title
Machine Learning, Virtual Reality, and Biomechanical Simulation to Aid Physical Rehabilitation

Author
Powell, Michael Ora

Publication Date
2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
SANTA CRUZ

MACHINE LEARNING, VIRTUAL REALITY, AND BIOMECHANICAL SIMULATION TO AID PHYSICAL REHABILITATION

A dissertation submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

COMPUTER ENGINEERING

by

Michael O. Powell

September 2021

The Dissertation of Michael O. Powell
is approved:

Professor Mircea Teodorescu, Chair

Professor Patrick Mantey

Professor Leila Takayama

Peter Biehl
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

List of Tables

Abstract

Machine Learning, Virtual Reality, and Biomechanical Simulation to Aid Physical Rehabilitation

by

Michael O. Powell

Physical rehabilitation is a continuum of healing for patients that extends beyond in-person clinical visits. Physical therapists expect their patients to continue performing stretches and exercises for weeks or months after the in-person treatment concludes. However, many patients fail to continue this practice due to the boring, repetitive nature of these exercises. In this dissertation, immersive virtual reality paired with biomechanical simulations was explored as a solution to this problem through the development of a rehabilitation game designed in collaboration with physical therapists. While participating in the National Science Foundation Innovation Corps Program with this work, we learned of another major problem facing physical therapists. Therapists are having difficulty accurately evaluating their patients on telehealth platforms. We addressed this problem by developing biomechanical simulations and machine learning models based on our rehabilitation game to propose new methods for remote patient evaluation. The findings of this work contribute to future physical therapy tool development that can aid therapists and help patients overcome accessibility barriers related to distance, time, travel, and disease that can prevent a patient from attending a clinic.

To my family, for always encouraging me to continue my education, your support through the many, many years, and your unending patience.

Acknowledgments

This work was collaborative in nature and would not have been possible without the support of many. I am extremely grateful for you all. My advisor, Mircea Teodorescu, taught me how to develop research based on existing skills to create unique opportunities that have made my education more fulfilling than I thought possible. Sri Kurniawan who shared her wealth of knowledge and expertise to help guide this project. Leila Takayama and Patrick Mantey whose expertise is indispensable in the review of this dissertation. reMarco Rolandi taught me what research is sensible in a graduate's time frame, what projects are rationale for an engineer, and opened possibilities to some of my long term goals.

Many peers have helped me throughout this endeavour. Thank you Sierra Catelani, for your continuous support, helping me navigate academic barriers, and your excellent edits to the many drafts of papers. Thanks to Aviv Elor for always diving into the next steps of our research without hesitation and for being the most reliable research partner a grad student can hope for. Special thanks to Ash Robbins, Maryam Tebyani, Gordon Kellar, Pavlo Vlastos, Sargis Yonin, Anca Popescu, and others I'm sorry that I missed for helping me develop my research ideas through white-boarding, discussion, and late-night hackathons.

Thank you to the Computer Engineering Department for this experience and education. Thanks to Ricardo Sanfelice and Jennifer Cornwell, who allowed me to work with them and show incoming students the possibilities within our department. Additionally, I'd like to thank Emily Gregg for helping me successfully navigate UCSC over the years. Without the support from UCSC with the Eugene Cota-Robles Fellowship, Campus Fellow Award, and National Science Foundation funding this research and my graduate education would not have been possible. I pursued the spirit of these awards, and I hope that this work proves to be a sound investment.

Chapter 1

Introduction

1.1 Motivation

Physical therapy and rehabilitation exercises are critical steps when recovering from acute bodily injury to muscles, tendons, or bones, as well as joint replacement surgery. Physical therapy can last weeks or even years, requiring a patient to see a therapist 2-3 per week. Additionally, patients are expected to continue doing progressive exercises on their own multiple times a week after in-person sessions are complete. This long duration and required diligence create many barriers for patients. Insurance only covers a limited number of sessions so anything beyond the patient has to pay, traveling to the clinic can be difficult or impossible for some populations, patients may have difficulty getting enough time off of work to visit their therapist as frequently as they are supposed to, and many patients don't continue their at-home exercises thereby limiting their recovery.

Physical therapists have attempted to solve these problems, but each solution has its pros and cons. Some therapists travel to their patients' home or work but commute time limits the number of patients they can see per day and drives up their price. Many therapists are moving to cash based clinics to avoid the issues involved with insurance. Insurance has limited reimbursement so clinics that accept insurance cut sessions shorter to see more patients per day to make up for lost revenue. By going cash based they can charge more per patient but provide longer sessions for a better quality of care. Another solution is telehealth and it has been widely adopted this past year with the pandemic. Telehealth helps cut out commute time, helps patients who can't travel, and is far more convenient. However, telehealth provides a very limited view of their patient making it difficult to perform evaluations. Therapists don't have the tools they need to get quantifiable metrics they need, and patients often have difficulty operating the telehealth platform and camera. A better solution is needed to help solve this problems patients and therapists are facing.

Virtual Reality (VR) has been used to help with various healthcare issues including treatment for post-traumatic stress disorder, balance improvement through exercise, alleviation of pain during occupation therapy, and for physical rehabilitation of older adults with disabilities. Head Mounted Displays (HMD) for virtual reality have improved greatly in the past two years with faster frame rates to reduce motion sickness, improved motion tracking for better immersion, and all-inclusive hardware which eliminates the requirement of a computer to operate the hardware. The entry price has dropped tremendously, a system can be purchased for under $300, making the technology accessible to many in the US. Virtual reality could be the key to solve many problems found in physical therapy.

This research project tested virtual reality media to determine which type of virtual reality system to utilize, developed a virtual reality game, and developed biomechanical simulations to determine in game metrics such as joint angles and torques. Afterwards, 130 physical therapists were interviewed to determine how our technology could be used to solve their greatest pain points. Two insights from these interviews were that telehealth is becoming widely adopted and that current telehealth platforms lack the tools required for accurate remote evaluation of patients' range of motion and strength. The next phase of the the research was to develop machine learning models to acquire biomechanical measurements that therapists need to evaluate patients that can be used at real-time. Our hope is to continue this project to develop a telehealth platform to improve remote evaluation and increase accessibility to rehabilitation care.

1.2 Applications

The motivation touches on the milestones of this work and we believe it can be applied to a several areas.

- Immersive Rehabilitation Game Design - We used biometric data and user feedback to aid in the design process of rehabilitation games along with input from physical therapists. Creating rehabilitative games requires collaborating with domain experts to ensure patient/participant safety. We believe this data and process can help other designers make informed decisions for their own games. These games can improve recovery and are often more appealing than repetitive exercises. Figure 1.1 shows our work being demonstrated on a television program about helpful emerging technologies for society.

Figure 1.1: Our work was featured on the Science Channel's show Crash Test World. Host Kari Byron is shown trying our one of rehabilitative games using the HTC Vive headset.

- Low Cost Motion Capture Systems - Typically motion capture systems are expensive and inaccessible to the general public. With our machine learning work we can provide joint kinematics and dynamics (a common use of motion capture systems) using only a low cost virtual reality system. This creates new opportunities for researchers and game designers who can not afford expensive motion capture systems.

- Remote Evaluation and Telehealth - This use case was touched on above. It has the potential to help many physical therapists and those going though a physical rehabilitation process. With accurate remote evaluation tools physical therapists will be able to monitor their patients' recovery and keep them safe from over extending or over loading their injured joints.

1.3 Contributions of this Work

This dissertation showcases our design and analysis to develop assistive tools for physical therapists. The contributions of this work are listed below.

- Determining which Immersive Virtual Reality (iVR) media would likely yield the best experience for users during exergaming.
- Designing a rehabilitation game in collaboration with physical therapists and testing over the course of two months with participants.
- Developing biomechanical simulations to determine joint kinematics and dynamics during rehabilitation game-play.
- Examining the novelty effect seen when participants try virtual reality for the first time and how their attitude toward use of new technology changes over time and with repeated use.
- Developing machine learning models to estimate biomechanical metrics during game-play for player monitoring and evaluation.

This dissertation concludes with a summary of insights, plans for making this work accessible to physical therapists and their patients, and recommendations for future research in this endeavour.

Chapter 2

Background

It's helpful to get a high-level view of this dissertation to understand why these topics are covered in the background. The main goal of this work was to create an accurate and low-cost tool so that physical therapists may properly perform remote evaluations of their patients to provide better rehabilitative care. Our in-lab research methods for obtaining biomechanical metrics is outlined in red in Figure 2.1. This method requires an expensive motion capture system, one example is OptiTrack, that can be as much as $50,000. The data from this motion capture system can be used as input into a biomechanical simulation software package, one example is OpenSim. These software packages often require domain expertise and programming experience to develop musculoskeletal models. Additionally, the run time for these simulations can be eight to twenty-four hours to run a single simulation. It is unlikely that a physical therapist would use this method as it is too expensive and takes too much time to process. The goal then became to make these biomechanical outputs affordable and to be determined in real-time. It turns out that virtual reality is a great technology to help with his endeavour. Figure 2.1 shows this low cost method outlined in green.

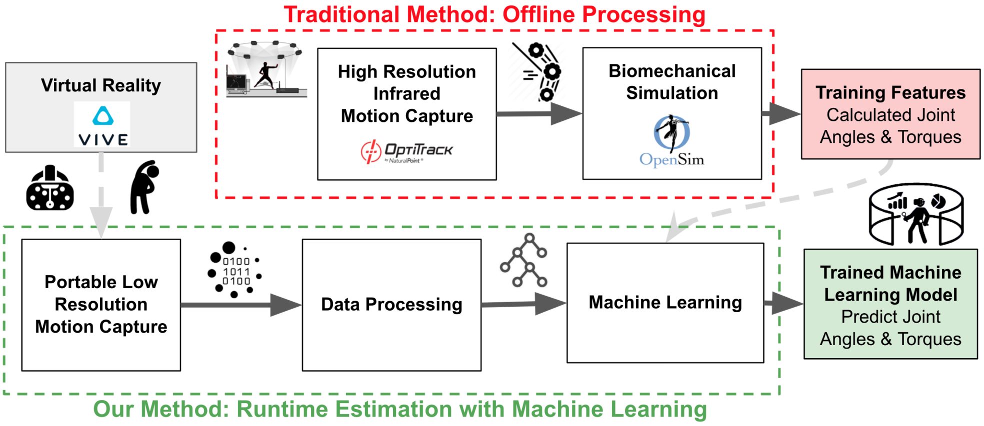

Figure 2.1: High level view of main research objectives. The red portion shows our in-lab research methods for obtaining joint angles and torques. The green portion shows our low-cost method to obtain the same metrics.

This research project requires piecing together current motion capture technology, biomechanical simulation software, head mounted display virtual reality systems, and machine learning methods. The important question to ask is "why now?" Only recently have all these technology pieces reached a point where the necessary data can be collected and processed and at an extremely affordable price. The background section outlines these technologies and why they are important now.

2.1 Virtual Reality for Health Applications

VR offers the unique ability to simulate complex situations that are critical to producing immersive experiences and is auspicious for improving psychologically-based health applications [4]. The use of VR intervention has been reported to achieve pain relief effects compared to an analgesic during wound treatment [5, 6]. Additionally, VR has been shown to help with Post Traumatic Stress Disorder [7, 8], Borderline Personality Disorder [9], Schizophrenia [10], and various phobias [4, 11, 12].

The multi-sensory, auditory, and visual feedback in a virtual environment can be crafted to persuade users further to comply with exercise protocols through increased directed stimuli [13]. Thus VR also holds immense potential in physiological rehabilitation as a useful tool for inducing task-based physical exercises [14]. The capabilities of multi-sensory real-time feedback have shown significant outcomes to achieve compliance with exercise protocols [13]. Numerous studies have displayed motor improvement in physiotherapy compared to traditional therapeutic intervention [13, 15, 16, 17, 18]. The biggest challenges of these studies were found to be technological constraints such as cost, inaccurate motion capture, non-user friendly systems, and a lack of accessibility [15, 19, 20]. Thus, there is a need to revisit this examination of VR for health with modern immersive Virtual Reality (iVR) systems [21].

2.2 Immersive Virtual Reality for Rehabilitation Applications

Immersive virtual environments can engage users and motivate them to overcome difficulty using virtual task goals in the context of rehabilitation. This leads to positive effects such as reduced discomfort, increased compliance, and flexibility [6, 22, 23, 24]. Moreover, the past five years have seen explosive growth in the field of iVR systems, stemming from a projected 200 million head-mounted displays systems sold on the consumer market since 2016 [21]. This mass adoption has been in part due to a decrease in hardware cost and a corresponding increase in ease of usability. To maximize immersion, iVR HMDs may be a promising tool that can fully engross the user in a virtual world.

Other researchers, e.g., Lindner et al., demonstrated the efficacy of therapist-guided psychotherapy through a low-cost iVR HMD system [25]. The authors found that the use of iVR devices successfully provided practical benefits for self-led and therapist-led intervention [25]. In a review by Won et al., iVR was found to be promising in assisting with the management of acute and procedural pain for adolescent patients by the process of neuromodulation through stimulating experiences [26]. In another survey, Li et al. explored and demonstrated the benefits of iVR applied to rehabilitation, disability management, surgical training, psychological disease treatment, and analgesic modality [27]. In Laver et al.'s review for VR therapy with stroke survivors, non-immersive VR therapy has been shown to improve arm function and activities of daily living for stroke survivors despite being less effective than conventional therapy [28]. Laver et al. also concluded that researchers designing VR rehabilitation programs should conduct pilot studies to evaluate usability and validity of the system and evaluate user's motivation, engagement and enjoyment.

The success of iVR therapeutic intervention is often attributed to the power of immersion, or the relationship between presence and emotion in an engaging experience [8]. Subsequently, a greater immersion corresponds to a better treatment response, and therefore, is beneficial to improving therapy experiences through virtual environments [29]. Providing engaging stimuli through immersive systems is a crucial factor for the player's experience [30]. The emotional response generated impacts user engagement and helps motivate players to continue with the objectives of the virtual experience [31]. Moreover, iVR systems provide a flexible environment for understanding player emotional response through the affordances of multimodal stimuli [32, 33]. Thus leveraging iVR stimuli to try to instigate a strong emotional response as done in psychotherapy may produce better results in exercise performance. Biofeedback devices may help us answer this question given that past studies had shown that biometrics can reliably record the response of users' emotional states [34].

2.3 Biofeedback and Physiological Response

Biofeedback devices have gained increasing popularity by using sensors to gather useful information about health states. For example, the impedance of the sweat glands, or Galvanic Skin Response (GSR), has been correlated to physiological arousal [35, 36]. This activity can be measured through readily available commercial GSR sensors, which have been used to measure arousal in media such as television, music, and gaming [37, 38]. Cameiro et al. analyzed non-immersive VR-based physical therapy that uses biofeedback to adapt to stroke patients based on the Yerkes-Dodson law [39], or the optimal relationship between task-based performance and arousal [40]. By combining Heart Rate (HR) with GSR, the authors examined gameplay by quantitatively measuring each user to determine where optimal performance is met. Another example can be seen with Liu et al, in which the authors were able to achieve 66% average emotion classification accuracy for users watching movies with only GSR sensors [41]. There is definite potential in evaluating the GSR and HR of each user to determine the intensity of the stimuli between different systems of VR.

GSR and HR are not the only biometric inputs that could be potentially leveraged into comparing immersive experiences. Commercially available Electroencephalography (EEG) sensors have shown great promise in capturing brain activity and even inferring emotional states [42]. Modern EEG sensors implemented through Brain-Computer Interfaces (BCIs) have successfully estimated user reaction to immersive stimuli during VR gameplay. In a review of over 280 BCI related articles, Al-Nafjan et al. examined how EEG-based emotion detection is experiencing significant growth due to advances in wireless EEG devices [43]. Accessible and low-cost BCIs are becoming widely available and accurate in emotion and intent recognition. These are being used for medical purposes as well as the non-medical domains of entertainment, education, and gaming [43]. In comparison with 12 other biofeedback experiments, studies that used EEG alone were able to reach 80% maximum emotion recognition [44]. Arguably, the most considerable challenges of BCI are costs, the accuracy of sensors, data transfer errors or inconsistency, and ease of use for devices [43].

Even with these challenges, EEG has been successfully used to understand conditions like Attention-Deficit/Hyperactivity Disorder (ADHD), Anxiety Disorders, Epilepsy, Autism, and Stroke [45, 46]. Brain signals that are characteristic of these conditions can be analyzed with EEG biofeedback to serve as a helpful diagnostic and training tool. For example, Lubar et al. used the measurement of brainwave frequency power during game events to extract information from reactions to repeated auditory stimuli. This provided the ability to perceive significant differences between ADHD and non-ADHD groups during this study[47]. By exploring different EEG sensors placements along a user's scalp and sampling multiple brainwave frequencies, different wavebands can be used to infer the emotional state and effect of audio-visual stimuli [48]. In another example, Ramirez et al. used the Alpha and Beta bands to infer arousal and valence, which are then respectively mapped to a two-dimensional emotion estimation model [42]. From these works, we concluded that there is the potential to analyze brainwaves during iVR stimulus to infer users' emotional responses.

2.4 Brainwave monitoring

Hans Berger, a founding father of EEG, was one of the first to analyze the frequency bands of brain activity and correlate it to human function [49, 50]. These wavebands have been extensively researched throughout the past eighty years, and while there are mixed opinions, we hope to use past research to contextualize brain activity during iVR exercise. Specifically, we want to understand the change from resting-state of the Alpha, Beta, Delta, Theta, and Gamma brainwaves induced by the gameplay.

The Alpha Band (Stress [51]) has been found to occur at frequencies between 7 to 12 Hertz and is generally associated with a neural activity relating to stress and conversely relaxation. Alpha activity is reduced with open eyes, drowsiness, and sleep [51]. **The Beta Band (Focus [52, 53])** occurs at frequencies between 12 to 30 Hertz, and is generally associated with focus, as well as active cognition such as arousal, anxiety, excitement, and concentration [52]. Increases in Beta waves have been correlated to active, busy, or anxious thinking and concentration [53]. **The Delta Band (Awareness [54, 55, 56, 57])** occurs at frequencies between 0.5 to 4 Hertz and is suggested to relate to awareness and sleep [54]. Delta waves have been found to have the highest activity during deep sleep, where the deeper the sleep, the higher the activity [55]. Researchers have also reported that this frequency band relates to memory interaction [56], such as flashbacks and dreaming [57]. **The Theta Band (Sensorimotor Processing [58, 59, 60, 61])** occurs at frequencies between 4 to 7 Hertz and is associated with sensorimotor processing [58]. This includes spikes in Theta activity for planning motor behavior [59], path spatialization [60], memory, and learning [61]. **The Gamma Band (Cognition [62, 63, 64, 65, 66, 67])** occurs at frequencies between 30 to 100 Hertz and has been correlated to thought, consciousness, and meditation [62]. Research has theorized Gamma activity is relational to conscious perception [63]. By studying meditation and mindfulness training, Gamma activity appeared elevated when a "conscious experience" would occur, such as shifting mental states in meditation [67]. There are mixed opinions on whether Gamma bands are reliable due to biological artifacts such as eye movement and jaw clenches [64, 65, 66]. However, many researchers argue that Gamma bands show evidence of correlating perception with careful signal processing [67]. Through combining active EEG sensing with VR gameplay, it may be plausible that the success of the VR stimuli in the virtual experience could be quantitatively measured.

2.5 Physical Therapy and Telehealth

In the United States, there are over 250,000 physical therapists, and this number is expected to grow by 47,000 in the next eight years to meet the growing needs of patients [68]. Telehealth plays an essential role in this growth by connecting patients to therapists and making care more equitable by helping patients overcome obstacles related to geography, time, finances, and access to technology [69]. Moreover, telehealth was determined to be effective in musculoskeletal practices having demonstrated outcomes and patient satisfaction comparable to in-person care [70]. Cottrell and Russell outlined considerations to apply when selecting a video conferencing telehealth platform for physiotherapy, which includes: appropriate privacy and security features, easy usability, clinician control of session beginning, financial cost, interoperability, the number of connections per session, and additionally built-in features (such as measurement tools, scheduling, playback, libraries, and questionnaires)[71].

2.5.1 Quantifying Telerehabilitation

During in-person and telehealth sessions, objective assessments that are valid and reliable are a crucial component to diagnose and treat patients [72]. Some standard evaluations during an in-person session include palpating a patient's affected injury, measuring Range of Motion (ROM) with a goniometer, determining strength using a resistive force (manual resistance, bands, or weights), mobility through a timed "Get Up and Go", and balance using the Berg's balance test. For the purpose of our research, we chose to work on objective tools related to ROM and joint forces that would aid therapists in their evaluations and monitoring of patients. Possible methods for measuring or estimating joint angles during videoconference telehealth sessions include digital goniometers, motion capture systems, computer vision applications, and sensor fusion techniques using inertial measurement units. However, for our iVR approach discussed in chapter 3, we aimed to use off-the-shelf systems (only controllers and headset) so there is no need for additional equipment. This provides a unique challenge as an iVR system itself does not provide joint angles; it only provides the position and rotation of each hand (from the controllers) and the head (from the headset).

2.6 Infrared Motion Tracking and Musculoskeletal Simulation

Motion capture is the most common practice for tracking the skeletal system's movements for biomechanical analysis. Reflective markers are placed on a subject which multiple cameras can track. With multiple views by cameras, the position of each marker can be determined and used for musculoskeletal simulation to determine body metrics such as joint angles or accelerations during movements. Within the DANSER lab we have an 8 camera OptiTrack system that uses the Prime 13W cameras (2.7 inches by 2.7 inches by 2.1 inches) and the accompanying motion capture software, Motive. This system is set up in conjunction with our CAVE (projection-based virtual reality display) that is 9ft by 9ft. Figure 2.2 shows one of the 13W cameras, the CAVE with the 13W cameras, and a player wearing the reflective markers that the cameras track. Within this area the cameras can estimate the markers with an error of approximately 0.5mm and record at a frequency up to 240 FPS. This data can then be used as input for biomechanical simulations.

Dynamic simulations can aid in analyzing performance as well as estimating the internal loading of the musculoskeletal system [2]. These simulations are extremely valuable in the context of rehabilitation and health. It is critical to find a balance for efficient exercise and speed of recovery, as overexerting strength and ROM may injure muscles. Finding this balance can be assisted by the use of modeling software such as OpenSim.

OpenSim is an open-source software system for developing musculoskeletal models and creating dynamic simulations of various movements [2]. The goal of OpenSim is to build a freely available library of movement simulations for the biomechanics community that has been validated and is ready for treating movement pathologies. The capabilities of OpenSim are vast and have been used to understand many applications such as human gait [73, 74, 75], design of assistive devices [76, 77, 78, 79], characterization of injuries [80, 81], and animal movement analysis [82, 83]. Gait mechanics have been well explored with OpenSim, but as of now, upper-body contributions are limited.

For our research, we desired to contribute multiple analysis techniques of various shoulder movements by utilizing the upper extremity model developed by Delp et al.

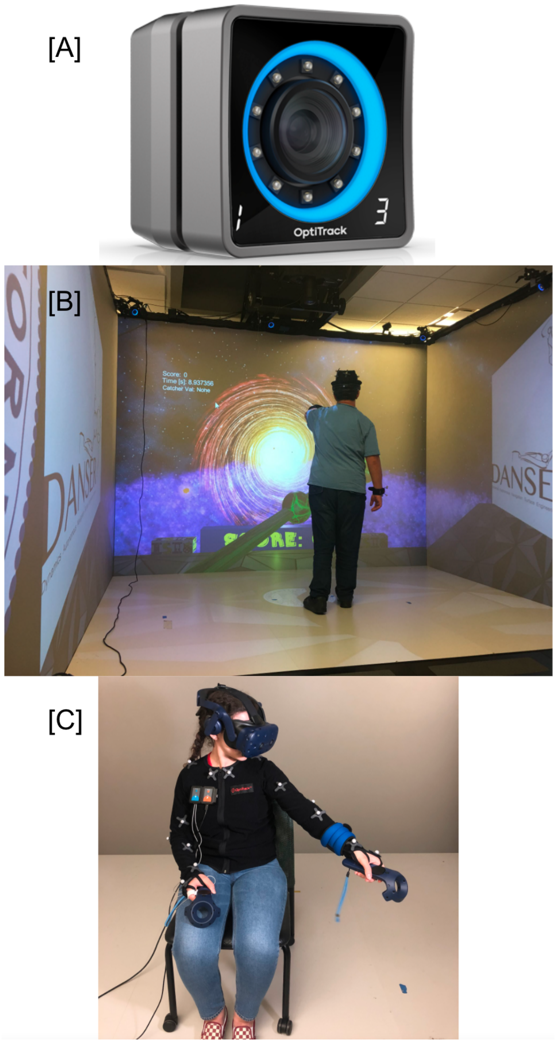

Figure 2.2: [A] shows one of the eight Optitrack 13W cameras located within our lab. [B] shows the game-play area within the CAVE and four of the 13W cameras are shown around the top edge of the screens (the blue rings). [C] shows a player playing a virtual reality game in the game-play area with several silver reflective markers attached to her upper body that the 13W cameras track.

[2]. We chose this specific model, as it includes all of the large muscle groups and the full ROM of the shoulder and elbow [2]. This simulation may prove valuable as it provides shoulder joint torques that can be tracked over an extended period. Torque is important because it describes the movement and force produced by the muscles surrounding the

joint [84, 85, 86, 87]. Prior research has examined the torque of upper-body exercise for more in-depth injury assessment; for example, Perrin et al. demonstrated that bilateral torque enables clinicians to more accurately set guidelines in the rehabilitation of varying athletic groups [88]. Another metric to consider is angular momentum [89]; this provides a metric to monitor user movement performance over several exercises, ensuring safety and preventing overuse. Several other studies have explored the benefits of quantifying angular momentum for robotic assistance [90], the severity of lower body gait impairment [91, 92], and how it contributes to whole-body muscle movement [93]. By examining average torque and angular momentum per session, we illustrate each user's average forces and amount of movement during gameplay as discussed in chapter four.

2.7 Machine Learning for Motion Tracking

Machine Learning (ML) is a method of data analysis that uses one or more algorithms to build a predictive model to estimate outcomes with new unseen input data [94]. There are many types of ML algorithms available that each utilizes different types of data and predictions methods. Typically these algorithms perform regression, clustering, visualization, or classification and can use probabilistic methods, rule-based learners, linear models (e.g., neural networks or support vector machines), decision trees, instance-based learners, or a combination of these [95, 96]. There are pros and cons to each and there is no universal best method for all data sets [97]. Instead, the type of input data needs to be taken into consideration, determine what type of prediction you want (e.g. binary classification, mulitclass classification, regression, ect.), identify the available types of models, and finally consider the pros and cons of those models. Some elements to consider with models are accuracy, interpret-ability, complexity, scalability, time to train and test, prediction time after training, and generalizability (the model produces acceptable results with unseen data) [98, 99, 100, 101, 102].

In our research, we use motion capture data to predict the output of biomechanical simulations (joint angles and torques). This means we have a supervised multiple regression task since our input and output data is already known, numeric and there are multiple input variables. Linear regression and decision trees are commonly used algorithms for these types of tasks. A decision tree is a simple flow chart structure made of nodes, branches, and leaves to go from an observation and traverse the corresponding nodes and branches to end up at a conclusion(also known as a leaf) [103, 104, 105]. This is a straight forward predictive model that has evolved in the ML community through many iterative steps and with each iteration results are typically improved. First, there is Bagging where multiple base learners (e.g. decision trees) with varying structure are built that can come to different predictions but in the end majority voting is used for the final conclusion [106]. The next step was Random Forest, where only a random subset of input features is used to build a collection of decision trees that then uses Bagging [107]. Boosting was the next evolution where weak models (e.g. regression or shallow decision trees) are built sequentially so that each new learner gives more weight to data that has been poorly predicted by the previous learners [108]. This method weights the data differently for each model so that each model performs better with that specific data compared to the rest of the models and in the end the weighted predictions are combined for the final output. Gradient Boosting came next and instead of adapting weights of the data, the algorithm tries to optimize an arbitrary differentiable loss function so that as sequential trees are added to the existing trees the loss is reduced (following the gradient) [109]. And finally, Extreme Gradient Boosting (XGBoost) builds upon all of these methods.

With the progression of computational power, memory, and cloud computing over the last decade, the number of machine learning applications continues to soar. With more groups looking to learn from ever-growing data sets, experts are looking to improve machine learning algorithm performance. One new algorithm to come out in the last five years is XGBoost. XGBoost was created at the University of Washington and has been one of the most widely used machine learning algorithms since being presented at a conference in 2016 due to its speed and performance [110]. XGBoost is a scalable decision tree ensemble algorithm that uses a gradient boosting framework but improves through system optimization and algorithmic enhancements [111]. For system optimization, XGBoost uses parallelization, tree pruning, and has cache awareness to help with training speed. Parallelization creates multiple branches of a tree in parallel to build tress quickly [112]. Rather than evaluating regularization at each node, the entire tree is built to max depth then walking from the bottom up determine whether each node and child are valid. This saves time and computation compared to top-down tree pruning [113, 114]. Cache awareness is used by buffering gradient statistics into each thread making efficient use of hardware resources [115]. The algorithmic enhancements include regularization (prevents overfitting by penalizing more complex models using LASSO and Ridge regularization), sparsity awareness (learns best missing values depending on training loss), weighted quantile sketch (finds optimal split point among weighted dataset using weighted quantile sketch algorithm), and cross-validation is built in at each iteration [110, 111]. For our research, we used XGBoost to predict biomechanical simulation outputs from iVR motion capture. This will be further discussed in chapter six.

Chapter 3

Comparing Virtual Reality Media

3.1 Introduction

There are many types of VR systems that fall into the categories of fully immersive (headsets or room-scale systems), semi-immersive (typically a mixture of physical and virtual environments such as a cockpit with screens instead of windows for simulating flying), and non-immersive (laptops and other two dimensional screens). Before designing and building a physical rehabilitation game, we wanted to determine which VR system would give players the best experience for recovery. Examining other studies, we determined that a fully immersive game would likely yield the best experience, but we still needed to decide if a room-scale system or a head-display system would be better. The study aimed to compare user experience between these two systems to help guide our future development. Our lab often works with users possessing cognitive disabilities to create rehabilitative experiences focused on their capabilities.

To ensure that the results are generalizable for users of varying cognitive abilities, we recruited forty users with and without cognitive disabilities, recognizing that both the iVR environments and games can be cognitively challenging for some users.

This study compared the Mechdyne Flex CAVE and the HTC Vive Pro 2018 HMD. We utilized an in-house customizable iVR exercise game that rewards users with and without a disability to overcome difficulties in exercising the weaker side of their upper body. We record each user's game behavior, physical movement, biosignals, and subjective response of gameplay and system use during gameplay. Through the differences in immersive experience between these two mediums, we aim to understand the effects of room-scale versus HMD based physical exercise.

Specifically, the goals of this study were:

1. To compare gameplay effects of the immersive exercise experience between the room-scale and HMD iVR mediums with natural arm movements.

2. To identify insights in system usability for users with varying cognitive abilities.

3. To examine the feasibility of the two iVR systems for exercise and healthcare.

3.2 System and Experimental Design

This study uses Project Star Catcher (PSC) [23, 22], an iVR experience designed to encourage upper-extremity physical exercises through motivating users to catch shooting stars in a cosmic virtual environment with their weaker arm. PSC uses a customizable mix of auditory, visual, and haptic stimuli as score incentives to motivate users to exercise. The game requires users to follow different arm positions and vary the range of motion in order to succeed in a star catch. The user receives three times as many points when using their weaker, weight constrained, non-dominant arm, but may also use their strong arm for fewer points. To perform well in the game, the user must use a large amount of full-body movement, including side stepping and reaching in many directions, and should comply with weak arm usage. Adults with developmental disabilities previously tested PSC. Our prior study showed that these users were able to understand and achieve the objectives of the game [23]. To ensure that the participants were challenged and understood the rule of the games, a weighted arm strap was utilized to examine weak arm compliance with the protocol from our previous exploration of PSC [23].

In this study, users were recorded playing PSC with both systems: EEG, GSR, and HR were collected at runtime as well as post gameplay surveys, as seen in Figure 3.1. The order of which system was played was counterbalanced (some users were tasked to the HMD first, and some to CAVE first) to prevent bias. We carefully designed the experiment so that users were exposed to a similar level of difficulty in both systems and similar features (e.g., soundtrack and screen brightness). In the CAVE, four walls are used to project multiple views at 90-degree offsets, whereas the HTC Vive implemented the native SteamVR camera allowing for a 360° view. From the viewpoint of user behaviors, the HMDs and the CAVE have many similarities; however, they are quite different in the level of immersion (i.e., users can still notice the outside world with the CAVE, while in the HMDs, they are completely isolated from the external visual stimuli). Additionally, the Vive HMD has more weight compared to the CAVE's motion capture markers.

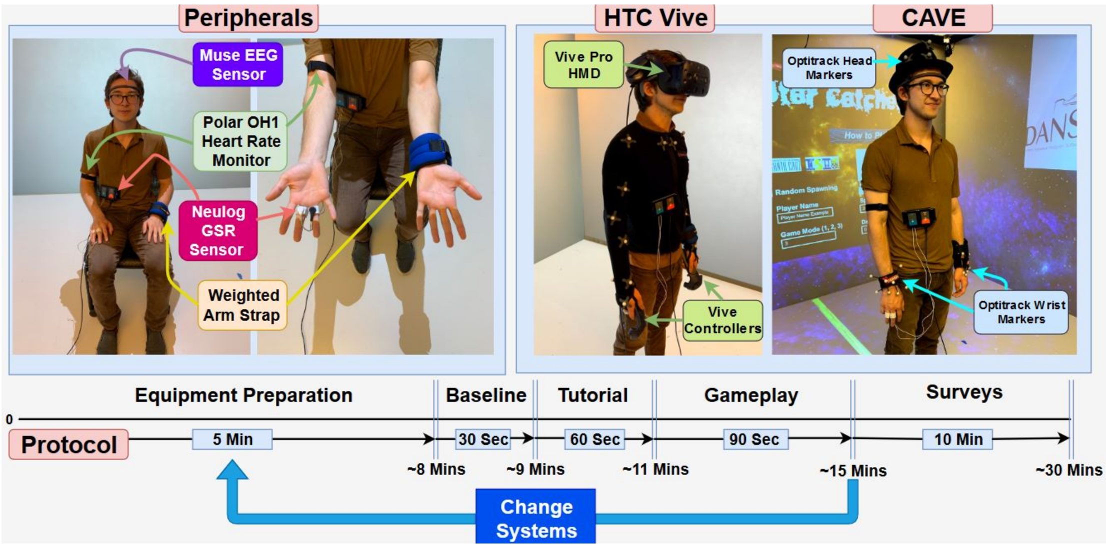

Figure 3.1: System Diagram and Experimental Protocol: Sensor placement (top left), Systems (top right), and experimental protocol (bottom).

3.2.1 Participants

Our participant cohort includes a mix of adults with Developmental Disabilities (DD) and college students. This study was approved by the Institutional Review Board (IRB) from the University of California - Santa Cruz (UCSC) Office of Compliance and Research Administration. For our volunteers with DD, three female and ten male users (ages ranged from 20 to 30) were recruited from the Santa Cruz Hope Services Day-Center and provided consent that had been vetted by their medical caregiver as understandable. Hope Services is Silicon Valley's leading provider of services to people with DD, such as intellectual disabilities, cerebral palsy, epilepsy, autism, and Down syndrome [116]. While they vary in their medical diagnosis, they all have a minimum cognitive ability specified by Hope Services' medical professional as likely to be able to comprehend the experimental protocol. Due to HIPPA regulation, we were not provided information on their diagnoses and the severity of their conditions. However, this information was available to our Hope Services collaborators during recruitment and formed the basis of their selection as volunteers. We shared our initial experimental protocol and questionnaires with Hope Services. We adapted our study protocol to ensure that these users could accurately reflect their feedback and participate in gameplay between the CAVE and HTC Vive. Additionally, a caregiver was present during all trials to help explain the study, the game, and survey questions as well as monitor safety, comfort, and provide further feedback about the participants. These thirteen users were selected by Hope Services medical professionals to ensure that they could articulate opinions about system preference and gameplay experience.

We also recruited 27 college students without any visible disabilities (12 male and 15 female with ages ranging from 19 to 28), who also provided written consent to participate. These students were recruited by flyers, word of mouth, and emails sent to the student body at UCSC. Through this diverse group of study participants, we were able to gather a mixed set of data between the CAVE and HTC Vive systems for the same iVR exercise game.

3.2.2 Data Collection

The following data was collected during the study:

1. HR - Polar OH1 Sensor [117]: an armband with an embedded optical sensor was utilized to wirelessly collect beats per minute by sampling HR activity at 1Hz.

2. GSR - Neulog GSR sensor NUL-217 [118]: a USB-200 logger sensor module was used to measure GSR at 5Hz sampling rate in micro-Siemens by attaching Velcro strap electrode points to the skin between the index and ring finger.

3. EEG - InteraXon Muse 2 - Brain Sensing Headband [119]: collects filtered brain-wave data of the prefrontal cortex. The application that communicates with the device uses a Cooley-Tukey FFT to extract waveband power from brain activity [120]. While this headset is relatively low resolution compared to other clinical-grade EEG systems, researchers have used Muse in understanding mindfulness [121], mental states [122], and event potentials [123].

4. iVR - The CAVE (Mechdyne Flex 1) and HTC Vive Pro 2018 systems. The room-scale CAVE system and HMD HTC Vive Pro implemented the Unity Game Engine to run the same iVR experience through PSC. PSC collected player data at 90Hz of motion capture pose and game behavioral events such as star catches [23]. Motion tracking was achieved with the CAVE through Natural Point Optitrack Motion Capture System [124], while the HTC Vive utilized its native lighthouse localization system for outside-in tracking [125].

5. Questionnaires - Modified Jennett et al. survey [126]: users completed the survey about user experience twice, once for each system. The users also completed a third survey that compared their preference of the two systems.

HR, GSR, and EEG measurements were chosen to give further insight into users' physiological responses to the gaming environment and provide quantifiable data beyond the game performance. The sampling frequency between all the sensors was not equivalent. As a result, data were exported to Comma Separated Values (CSV) files post hoc and synchronized for each baseline and gameplay using Python scripts. A custom MATLAB script was implemented to collect all sensor data in a single nested struct for comparison, while also running raw sensor data through a smooth-data moving window filter. Statistical significance between systems was determined in MATLAB through Wilcoxon signed-rank tests, which is a non-parametric statistical method to compare two related groups by mean rank difference [127, 128].

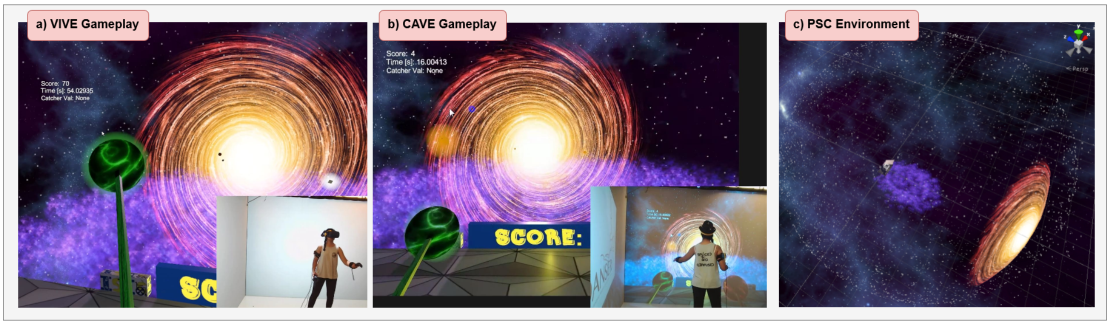

Figure 3.2: System Gameplay: a) A user catches a shooting star with the HTC Vive. b) A user prepares to catch a shooting star with the CAVE. c) The PSC virtual environment.

3.2.3 Experiment Design

Our experimental protocol consisted of four stages that were completed one time on each system, followed by a final set of surveys. This order can be seen in Figure 3.1 and is described in detail below:

1. Equipment Preparation: The HR monitor was placed on the dominant arm, and two GSR electrodes were positioned on the middle two fingers for the participant's dominant hand. The EEG sensors was set on the forehead located on the AF1, AF7, TP9, and TP10 prefrontal cortex positions. A weighted arm strap (selected to be approximately 3% of the participant's body mass) was fixed to their non-dominant wrist to challenge and remind the user to catch stars with their non-dominant arm. Finally, either the HTC Vive controllers or Optitrack markers for the CAVE were given to the user depending on the counterbalanced system starting order.

2. Baseline: Before any gameplay, the participant was asked to stand still with their arms at their side and eyes closed for 15 seconds, followed by 15 seconds with their eyes open. We recorded sensor data during this step in order to determine changes from resting-state to gameplay.

3. Tutorial: The evaluator then started the tutorial game, began to give scripted verbal instructions on how to play, and answered any participant questions. The evaluator administered the tutorial for approximately 60 seconds to ensure the user had grasped the concepts of the game.

4. Gameplay: After the tutorial, the evaluator set up the game, let the participant know they had 90 seconds to play the game, would no longer receive feedback or verbal instructions, gave a count down, and began the game. After the 90 seconds of gameplay, the evaluator gave a verbal countdown to warn the participant the

test was ending.

5. Change Systems and repeat (i)-(iv): Next, the participant was outfitted with the other game system. There was another baseline measurement, tutorial stage, and gameplay identical to the previous ones.

6. Surveys: The evaluator then removed the game system and provided a chair for the user to sit while filling out the surveys.

Between each stage was a transition period of about 1-3 minutes of rest time. A table comparing baseline biometric state indicated that this rest period was adequate with no significant differences of biometric measurements between recordings, as shown in Table 3.1.

3.3 Results

Session data was post-processed using the Mathworks MATLAB 2018b Statistics and Machine Learning Toolbox [129]. We examined each of the user's recorded metrics between systems and groups for box-plot distribution, significance, and similar metrics. Significance was determined through a Wilcoxon signed-rank test, a confidence statistic used to compare non-parametric data such as the samples obtained in our study [130]. The intent of this data collection was to determine the physical and biometric performance between each system and user group in the context of feasibility, immersion, and potential for iVR exercise experiences. These results indicate that both systems are useful in obtaining high levels of compliance with game goals during physical exercise. We define compliance as the rate of catches with the weighted non-dominant arm over the total amount of catches. From these metrics, the HTC Vive was found to be significantly more effective than the CAVE in inducing more significant movement of the non-dominant limb, a greater resting-state change in biometric response, a more significant emotional response, and an increased immersion. These findings are particularly exciting as prior studies that have explored CAVE and HMDs have not found significance in their task-based comparisons [131, 132, 133, 134, 135]. We discuss these findings in the following subsections.

3.3.1 Physical Movement and Gameplay

Recording runtime motion capture and behavioral game datum served to help understand the physical performance of the users across different cognitive abilities between the two iVR systems. As we are interested in how the users with and without cognitive impairments differ in their gameplay behaviors, we separated the users into two groups. Physical displacement of each user's non-dominant arm, dominant arm, and head positions are shown in Figure 3.3. For both user groups, the HTC Vive induced significantly more gameplay movement of all tracked limbs when compared to the CAVE. The user group with disabilities also had more movements when they were using HTC Vive than the cohort without disabilities.

To examine compliance, we set the game mechanics so that users achieve higher scores when performing successful catches with the non-dominant arm than when using the dominant arm. We define compliance as the total catches with non-dominant weight-constrained arm over total star catches. This can be seen in Figure 3.4 along with a successful star catch rate and game score. Both user groups had a significantly higher catch rate (successful movement completion) on HTC Vive, yet did not hold significant differences in compliance between the two systems. The groups differed in game scores, where users without disabilities had a significantly higher score with HTC Vive than with the CAVE, but users with disabilities do not have significant differences in scores between the two systems.

PSC varies the difficulty of star catches by movement speed through spawning bronze, silver, and gold stars as slow, medium, and fast respectively. For example, bronze stars are the easiest to catch as they move three times slower than gold, but the reward is also three times less in score. Figure 3.5 highlights successful star catches in terms of difficulty. As expected, both groups completed significantly more easy and medium catches with HTC Vive. However, users with disabilities did not have a significant difference in hard gold catches between systems. The group without disabilities caught more stars than users with disabilities across all difficulties, which was expected.

To summarize, both groups performed significantly better on HTC Vive in terms of physical movements and successful star catches, yet groups differed in strategies where the users with disabilities did not significantly overcome challenges associated with hard catches between the two systems.

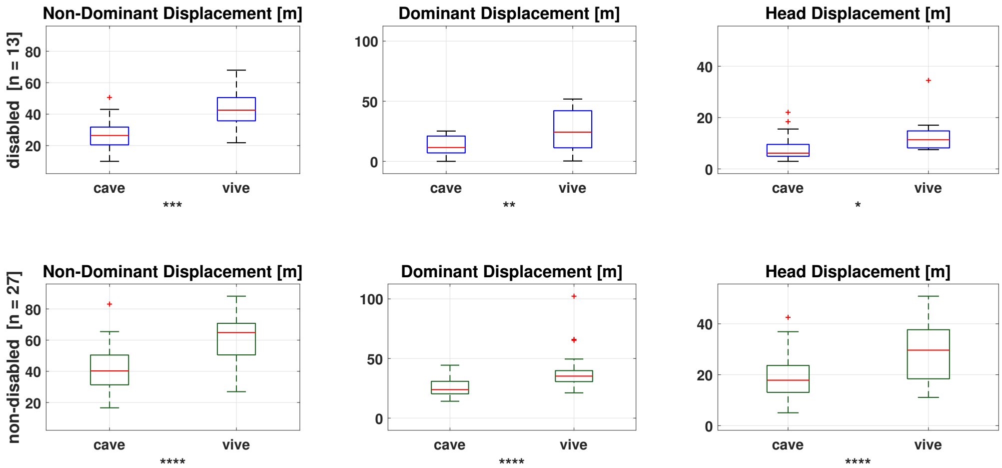

Figure 3.3: Player movement of users with disabilities (row one) and without disabilities (row two). Wilcoxon significance level between CAVE and VIVE is indicated in asterisk notation and “ns” indicates not significant (highlighted in red). Note that Non-Dominant Displacement indicates the total movement of the weighted arm during Project Star Catcher Gameplay between systems.

3.3.2 Biofeedback Responses

Three sets of biofeedback data were collected during and before gameplay to infer physiological activity: Heart Rate (HR) as a measurement of physical intensity, Galvanic Skin Response (GSR) as a marker of arousal, and brainwave activity (EEG) as inferences for stress (Alpha power), focus (Beta power), awareness (Delta power), motor activity (Theta power), and cognitive state (Gamma power). For the context of the study in this chapter, these physiological effects from biometric activity are used to contextualize resting-state change induced from gameplay between the two systems. A pre-gameplay baseline was recorded before every user trial to determine and normalize possible abnormalities produced from daily living – for example, if a user was overstimulated by an intense conversation before testing, this stimulation would be offset by

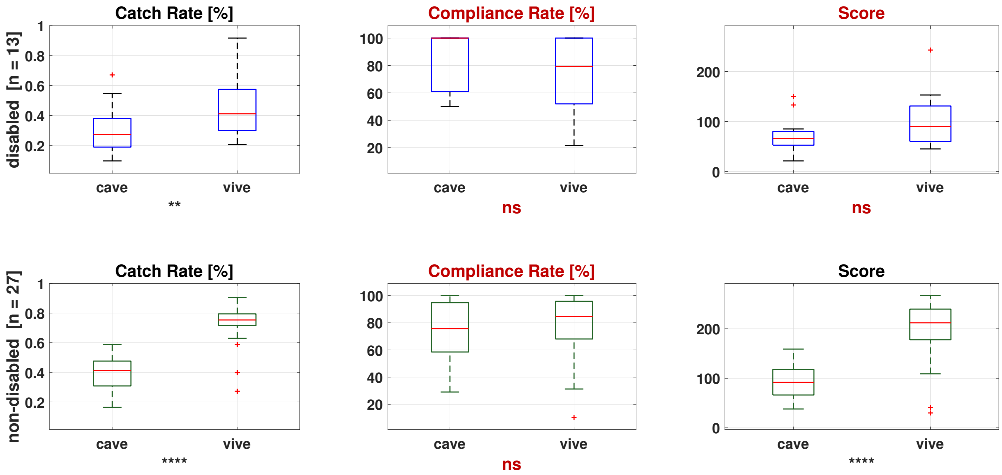

Figure 3.4: Gameplay score and success rates of users with disabilities (row one) and without disabilities (row two). Wilcoxon significance level between CAVE and VIVE is indicated in asterisk notation and “ns” indicates not significant (highlighted in red).

examining the difference in the gameplay and baseline states. We were careful to not unnecessarily converse with users during the study to avoid individual differences due to protocol deviation. The results showed that the HTC Vive produced considerably more biometric changes compared to the CAVE, with noticeable differences between the two user groups. The Wilcoxon significant levels between pre-gameplay states of both user groups are shown in Table 3.1, and indicate no significant difference between pre-gameplay states between systems, with the exception of the gamma band for the group without disabilities.

Figure 3.6 shows the resting-state change of HR and GSR induced by gameplay with PSC. Users without disabilities had significantly higher HR and GSR when using the HTC Vive, which may indicate higher intensity in physical activity and arousal. On the other hand, users with disabilities had no significant differences between the two

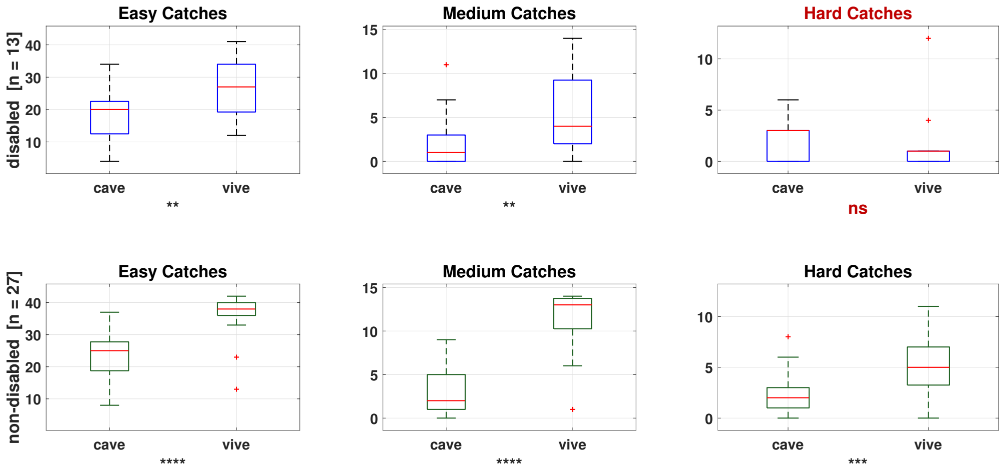

Figure 3.5: Successful star catches with difficulty of users with disabilities (row one) and without disabilities (row two). Wilcoxon significance level between CAVE and VIVE is indicated in asterisk notation and "ns" indicates not significant (highlighted in red).

systems, yet HR tended to remain at a definite increase from the resting-state baseline, and much of the GSR distribution for the CAVE indicated a decrease of arousal from resting-state. This may indicate that users with disabilities were either overstimulated before playing PSC with the CAVE or that the CAVE was ineffective in stimulating arousal for these users. Table 3.1 suggests similar pre-gameplay states, so it was more likely that the CAVE itself induced this negative change in arousal. For the cohort without disabilities, both systems produced an increase in all biometric recordings from resting-state, and the HTC Vive had a significantly higher increase of intensity and arousal than the CAVE from the HR and GSR readings. Interestingly, brainwave activity represented an inverse outcome between the two user groups.

The resting-state change of the different EEG brainwave bands induced by gameplay with PSC is displayed in Figure 3.7. Both user groups had significantly higher

Beta and Gamma power when using the HTC Vive against the CAVE, which may indicate an elevated level of focus and cognitive processing. The groups differed where the cohort with disabilities had significantly higher Alpha, Delta, and Theta (stress, awareness, and motor processing) power. Furthermore, the group with disabilities generally experienced negative resting-state change on CAVE for Alpha, Beta, Delta, and Theta bands, which may imply the users did not remain focused and lost awareness as well as a motor activity when compared to resting-state. This negative resting-state change is consistent with the change seen with CAVE for HR and GSR; however, all brainwave bands were significantly higher on HTC Vive inversely to the relationship seen between the two groups in Figure 3.6.

In general, these biometric recordings suggest that the HTC Vive induced higher focus and cognitive processing than the CAVE for both groups. Unlike the group without disabilities, users with disabilities had significant increases in all bands of brain activity. Conversely, the CAVE induced a lower power from resting-state change for the beta, delta, theta, and gamma bands, unlike the HTC Vive, which resulted in all significantly higher powers than resting-state. This differs from the group without disabilities, where all brain activity remained at a positive change regardless of the iVR system medium. This outcome is especially interesting as it may indicate that iVR system mediums have a more considerable effect on the mental state for adults with cognitive impairment. To further understand these results, each user was queried for subjective response in our immersion and system preference questionnaires – it is through this medium that we hope to reinforce and better understand the physical and

	with disabilities			without disabilities		
data type	CAVE mean (std)	VIVE mean (std)	sig	CAVE mean (std)	VIVE mean (std)	sig
HR [bpm]	104.9 (38.97)	116.6 (29.74)	ns	94.5 (24.53)	91.1 (21.11)	ns
GSR [uS]	2.49 (1.330)	2.54 (1.195)	ns	3.18 (2.322)	3.14 (2.123)	ns
Alpha [bels]	0.70 (0.396)	0.60 (0.253)	ns	0.66 (0.139)	0.63 (0.143)	ns
Beta [bels]	0.52 (0.361)	0.44 (0.331)	ns	0.50 (0.221)	0.39 (0.201)	ns
Delta [bels]	0.98 (0.503)	0.95 (0.403)	ns	0.75 (0.325)	0.75 (0.278)	ns
Theta [bels]	0.54 (0.418)	0.43 (0.283)	ns	0.41 (0.195)	0.40 (0.180)	ns
Gamma [bels]	0.34 (0.367)	0.17 (0.394)	ns	0.30 (0.323)	0.10 (0.279)	***

Table 3.1: Biometric baselines taken at resting-state between two user groups for both systems. "sig" indicates Wilcoxon significance level in asterisk significance notation, with "ns" indicating no significance. No significant difference was found between pre-gameplay states for all groups with the exception of the Gamma band for the non-disabled group.

biometric performance of our users.

3.3.3 Response for Immersion, Emotion, and System Preferences

In this study, we used two surveys to collect subjective responses between the HTC Vive and the CAVE from our two user groups. The immersion questionnaire was adapted from an extensively explored survey by Jennett et al, which measures immersion and presence in games [126]. For the group with cognitive disabilities, pre-experimental trials were run to understand the feasibility of the original immersion survey and help us modify the survey. These trials were useful as they provided us with some insights. Generally, users would lose interest in the high number of questions in the original survey. Additionally, the phrasing of most of the original questions was often too complicated for users to comprehend fully and required the Hope Services Caretakers to intervene and give further explanations and provide examples. Lastly, many of the users were not always able to communicate their responses verbally – usually giving a

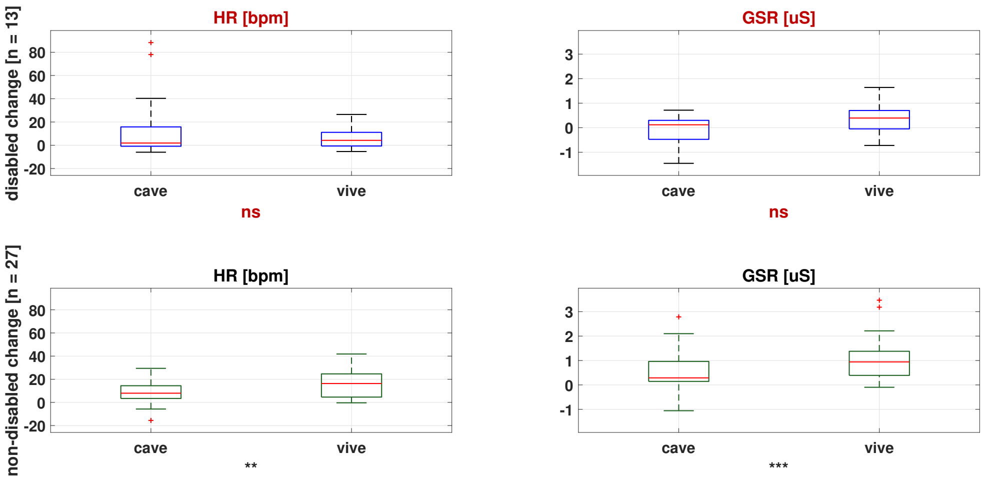

Figure 3.6: HR (in beats per minute) and GSR (in micro-Siemens) resting-state change from gameplay of users with disabilities (row one) and without disabilities (row two). Wilcoxon significance level between CAVE and VIVE is indicated in asterisk notation and "ns" indicates not significant (highlighted in red).

thumbs up, down, or sideways. With this trail-testing in mind, we condensed the Jennet et al. survey to ten simplified questions in collaboration with healthcare professionals. Furthermore, a checkbox emotion question and system preference survey were created to enable more significant user input from the group with disabilities. Our final version of the questionnaires consisted of one survey with ten immersion questions on a subjective scale, one question on intense emotions felt, and a second survey with three questions on system preference and a section for additional comments. Through this process, we were able to gather more significant input from both user groups for comparison with the biometric and game datum collected.

The immersion survey results, as seen in Figure 3.8, indicates response on statements querying presence (Q3-4 & Q8), engagement (Q1-2 & Q9-10), and effort (Q5-7) concerning gameplay between the two systems and groups. Questions Q8-10

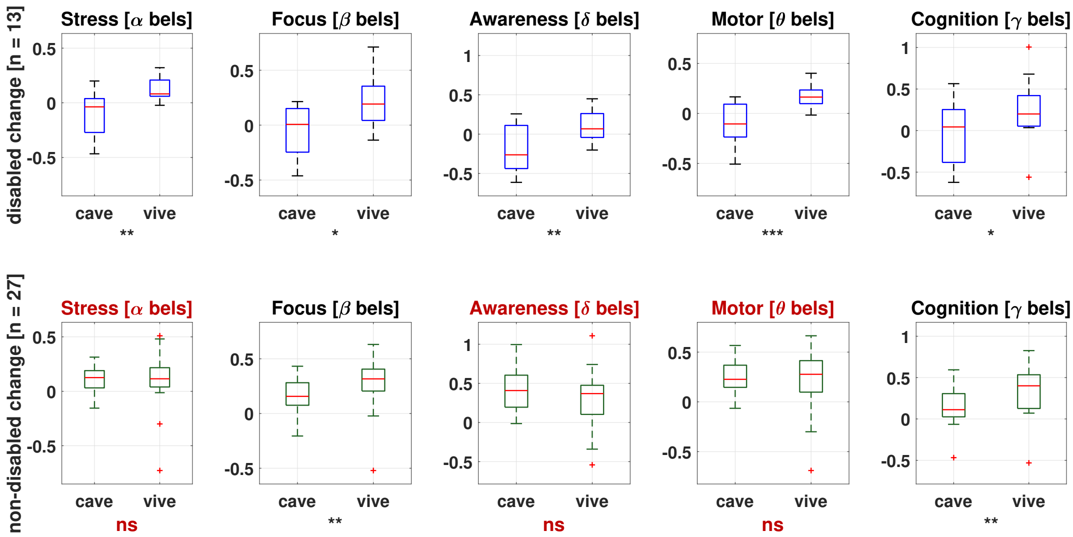

Figure 3.7: EEG brainwave power in bels from resting-state change induced during gameplay for users with disabilities (row one) and without disabilities (row two). Note that stress, focus, awareness, motor, and cognition are represented by the alpha (α), beta (β), delta (δ), theta (θ), and gamma (γ) band powers. Wilcoxon significance level between CAVE and VIVE is indicated in asterisk notation and "ns" indicates not significant (highlighted in red).

have reversed scales to ensure respondents read the survey carefully. A majority of users from both groups indicated that presence, engagement, and effort was higher on HTC Vive than CAVE. The groups differed in agreement, where higher percentages of users with disabilities felt they "lived in the game world," were distracted "from my real life," and "put a lot of effort into the game." Interestingly, the majority of the disabled cohort found the game to be not challenging (Q9), unlike the non-disabled cohort, even though their physical performance was, on average less than the non-disabled group (as seen in Figures 3.3 and 3.4). The disabled group responses to immersion questionnaires were nearly identical between all users regardless of system, which may indicate a lack of comprehension of the survey questions regardless of our modifications or that users generally responded positively to all survey questions. There are slight differences in the distribution, which may indicate that HTC Vive was received better in comparison to the CAVE (Q1, Q7 & Q10). This was not the case for the non-disabled cohort, where users had significantly higher agreement rates with the HTC Vive than the CAVE. These immersions were most likely influenced by the emotional response felt during gameplay.

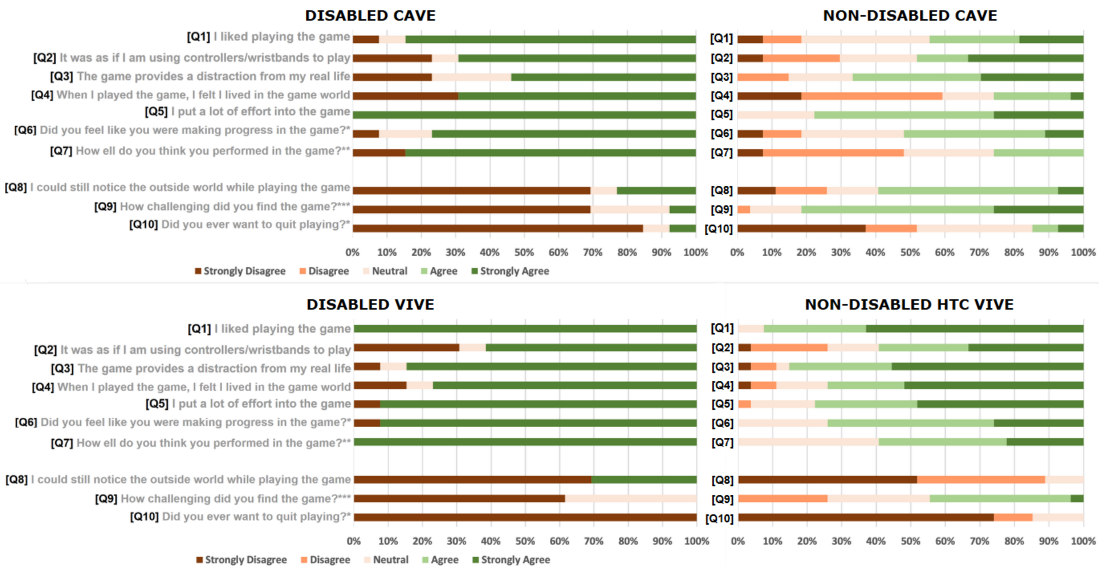

Figure 3.8: Subjective rating questionnaire responses for the between user groups and systems. For Q1-7, strongly agree is the desired outcome. For Q8-10, Strongly disagree is the desired outcome. Disabled user responses were modified to 3 point scale as recommended by healthcare professionals from Hope Services, CA, to increase accuracy. * = "Not at all" to "a lot," ** = "Very poor" to "very well," *** = "Not at all" to "very challenging."

Self-reported emotions felt during gameplay can be seen in Figure 3.9. At the end of each session, users were tasked with checking off any intense emotions they believed to have felt during the three minutes of playtime between the systems. The emotions cover a wide range of feelings from "Happy/Joyful" to "Neutral (no emotion)" to "Angry/Hateful/Disgust." For the purposes of visualization, such emotions are organized subjectively from top-down positive to negative in Figure 3.9. All users from each group reported feeling at least one intense emotion from gameplay between each system. The non-disabled group generally reported more feelings of positive emotion with the HTC Vive, and CAVE was shown to receive higher responses for negative emotions such as angry and embarrassed. Conversely, minimal difference between the two systems on self-reported emotion was found for the disabled group – where CAVE had a slightly higher emotion response rate than HTC Vive by one or two users. The ratio of intense feelings for users with disabilities was also significantly higher than their non-disabled cohorts, which may be in line with the increased disabled group brainwave activity, as seen in Figure 3.7. These near-identical distributions in emotions felt between the systems may indicate that the majority of users with disabilities may just be answering these surveys identically. This behavior, however, was not seen in the preference survey between systems for the group with disabilities.

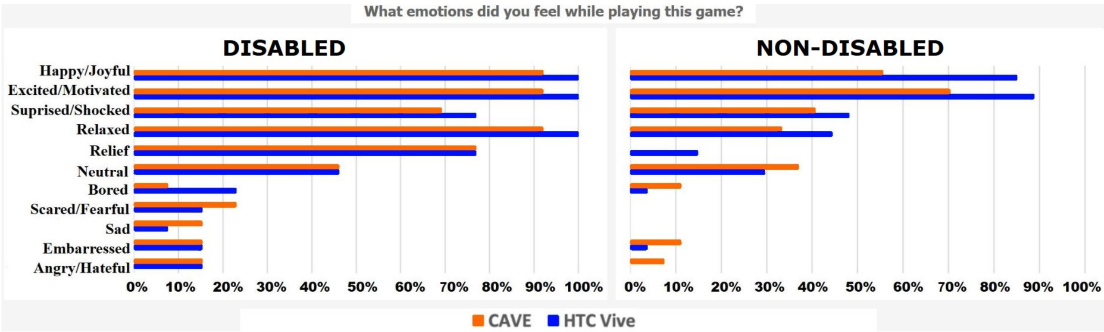

Figure 3.9: Self-reported emotions strongly felt between the two different systems and user groups.

At the end of the experiment and after the two immersion/emotion surveys, a preference survey was given to each user asking which system was preferred and why.

Users were also given the option to fill out checkboxes on indications such as comfort, ease of use, and engagement as well as an input field for additional comments. Figure 3.10 showcases these final preference results. The majority of both groups preferred HTC Vive over the CAVE; however, 100% of users without disabilities preferred HTC Vive, unlike the 62% of the cohort with disabilities. These groups appeared to generally differ in system preference by emotion and comfort with CAVE when compared to ease of use and immersion with HTC Vive. For the group with disabilities, the users who chose the CAVE indicated the most active reasoning was "it made me feel relief," "it was easier to use," and "it was more comfortable," whereas the users with disabilities who chose HTC Vive indicated top reasoning to be "it was funner to use," along with a near-identical indication of greater comfort, ease of use, immersion, and relief. The users without disabilities' top reasonings for unanimously choosing HTC Vive over CAVE was "it was easier to use," "it felt more immersive," and "it was funner to use."

Additionally, about 50% of participants wrote in or verbally addressed additional comments about system preference. A word cloud of these comments can be seen in Figure 3.10, where the largest words indicate the most reoccurring topics of discussion. Only four of the users with disabilities who preferred the CAVE left additional comments and indicated they enjoyed wearing the motion capture hat, unlike the HTC Vive HMD. For other users, comments were left about navigation, perception, latency, and freedom of movement appearing best on HTC Vive against the CAVE. Participants felt PSC was more stimulating on the HTC Vive than CAVE as the colors were crisper, the depth perception felt more viable, and the controls were more natural, according to them. One user even indicated the preference of blocking reality out with HTC Vive, unlike the CAVE, as that they felt the HTC Vive was "more immersive [because] my virtual self was already in there [the game]."

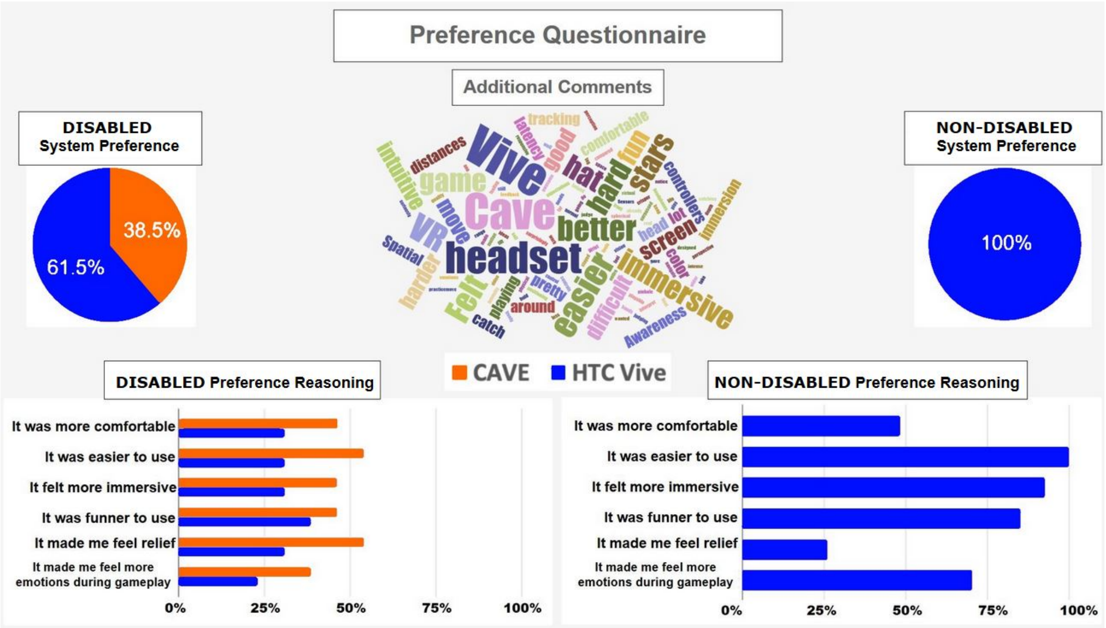

Figure 3.10: System preference between the two user groups with reasoning for preference.

To conclude, the HTC Vive is the preferred system between both user groups. The HMD based system was perceived to have a higher sense of immersion, ease of use, and enjoyment of gameplay than the room-scale alternative. These responses are in line with both the physical performance of each user group as well as the biometric response. The significantly higher brainwave activity among users without disabilities on HTC Vive is in line with the self-reported levels of immersion, where they saw significant increases in arousal and physical activity and subsequently unanimously preferred the HTC Vive. The responses of users without disabilities tended to be emotionally based, with a significantly higher distribution of these users' self-reported intense emotions felt. Lastly, users generally scored higher, moved more, and caught more stars on the HTC Vive against CAVE – this can be explained by users feeling the HTC Vive was "more fun."

3.4 Discussion

This study explored the experience of adults with varying cognitive abilities when using room-scale and HMD based iVR systems for gamified physical exercise. A mixture of motion capture, game behavior, and biofeedback, along with questionnaire data, was collected. The HTC Vive, a widely adopted commercial iVR HMD system, showed significant benefits of use when compared to the CAVE during physical exercise with the game we built. This section highlights our findings.

3.4.1 Key Findings Between the Two User Groups

We explored two user groups in this study: adults with cognitive disabilities and college students, in an attempt to make sure that our results and the design implications for our findings can be generalized across cognitive abilities. Through our testing of these systems, the data we collected helped us address our study goals to formulate the following interpretations:

- **Users with cognitive disability were more emotionally receptive to iVR exercise.** All bands of brainwaves were seen to be significantly different between the systems, which may have influenced the noticeably higher self-reported emotions by the group with disabilities. From the immersion survey, we saw a similar response for virtual presence, engagement, and effort given during gameplay. This may indicate that our game was a successful experience in inducing immersion regardless of the system. Furthermore, the preference survey indicated that it was, by a majority, guided by emotion. For the users who chose the CAVE over the HTC Vive, the top reason was due to feelings of relief. Additional comments indicated users enjoyed the way the motion capture hat felt and looked in comparison to the HTC Vive HMD. These users chose the CAVE even though they often overcame greater difficulty, caught more stars, and had higher movement with the HTC Vive. From an engagement and immersion perspective, designing future experiences for adults with cognitive disabilities may be improved by expanding on this emotive perspective. PSC uses score and sensory feedback as a motivator to keep the user engaged over their exercise sessions [23]. Another one of our previous studies has explored iVR games to strictly follow exercises by protecting a "cute" virtual butterfly in "Project Butterfly" (PBF) [136]. With therapeutic goals in mind, designing iVR experiences like PSC and PBF where the user is in complete control of the environment and is guided by emotive based incentives can be an excellent approach for iVR physical exercise experiences. The physical tasks require only user movements, and the objectives of the game are simple enough to start and interact with the environment without reading an instructional guide on controls or game objectives. As a result, emotive task-based iVR experiences for exergaming will most likely increase the adherence, engagement, and success of an exercise protocol.

- **Users without disabilities were more physiologically receptive with iVR exercise.** These users physically performed in all areas of gameplay with more significant movements, successful catches, difficulty overcome, higher game scores. Arousal and physical intensity were seen to be significantly higher on the HMD when compared to the room-scale medium, unlike the user cohort with disabilities. Resting-state change of brain activity was insignificant on three out of the five wavebands, where only beta and gamma were found to have a significant change with the HTC Vive. Subsequently, all 27 of these participants unanimously preferred the HTC Vive. Unlike the cohort with disabilities, whose preference was primarily driven by elements such as the feeling of the motion capture hat and other emotive reasoning, the group without disabilities valued the HMD's ease of use, control, and increased immersion for completion of exercise tasks. For these users, the higher physical performance with HTC Vive impacted immersion, emotion, and system preference, as shown by the apparent significant differences for HMD questionnaire responses. HTC Vive is the clear winner here when compared to CAVE.

- **Both cohorts performed better with HMD based iVR exercise.** Our user groups shared many similarities. The HMD system induced more significant physical movements, difficulty overcome, brainwave activity of the beta (focus related), and gamma (cognitive processing related) bands, and they both subjectively reported higher levels of immersion, engagement, and effort during gamified exercise with HTC Vive. The majority of both groups preferred the HTC Vive over the CAVE, even though the CAVE provided an untethered physical medium for experiencing the virtual world. HMDs enable a full virtual immersion where users preferred this medium because the experience felt "more immersive [because] my virtual self was already in there [the game]." This detachment from reality proved to result in higher engagement, which may have attributed to a better physical and biometric performative response. Based on our results, we can conclude that for future experiences employing gamification for task-based exercise goals, HMD based systems of iVR (which are significantly cheaper and more portable than the room-sized versions) are the apparent decision to maximize performance and engagement.

3.4.2 Has Modern Commercial HMD Based iVR Surpassed the Research Grade Room-Scale Medium in Healthcare Context?

This study has shown that HMD based VR is a better medium compared to the room size version for maximizing physical performance, immersion, engagement, and effort of task-based exercises. Research has shown that the full exclusion of the real world provided by the HMD enables higher immersion, which is a powerful tool to distract users from pain and discomfort [5, 6]. Applying these immersive effects to overcoming adversity and difficulty in exercise is useful. In a past study with stroke survivors, PSC has shown that the gamification with iVR of physical therapy can increase compliance by nearly 40% compared to traditional therapies with the HTC Vive [23]. Subsequently, from the perspective of accessibility, accuracy, and affordability of exercise-based healthcare, HMD based commercial iVR systems may have finally surpassed the alternative and more costly room-scale mediums.

The CAVE does not exclude reality from the virtual world, as the user's physical body itself becomes a part of the visual iVR experience. Researchers have argued that this nature of the CAVE – to be able to include multiple people in a space with their physical presence – can be advantageous for collaborative task-based needs [134]. However, we argue that modern commercial HMD based systems have fully surpassed the advantages the CAVE with regard to multi-user applications. For a fraction of the price, multiple HTC Vive-like HMD systems can be purchased and may enable multi-user interaction via virtual avatars from any location through the internet. The cost of an installation of a CAVE and its lack of mobility makes the CAVE less flexible compared to the HTC Vive. New inverse kinematic techniques are being developed and shared across the research community, with integration for mass-produced systems like the HTC Vive to show full-body motion capture approximation and ease of implementation [137]. Furthermore, there is a vested interest in producing full-body motion capture by industry competitors for the future of iVR interaction with HMDs [138, 139]. In summary, iVR HMDs are gaining popularity: the cost of headsets are decreasing, systems are becoming ever more mobile and untethered through new inside-out sensor fusion tracking techniques, and new input mediums such as hand tracking, eye tracking, voice control, and even more are being integrated [21]. We should also note that the integration of these features into a volumetric space like the CAVE would cause a significant increase in cost compared to the HMD medium.

3.4.3 Considerations and Limitations

This study was one of the first to compare room-scale and HMD based iVR through game performance, biofeedback, and questionnaires for exergaming with adults of different cognitive abilities. However, some limitations need to be considered.

Past studies with the PSC iVR framework have shown great potential to increase compliance for adults with physical disabilities [23, 136]. Nevertheless, these users have not been explored with CAVE in this study due to resource constraints. Future studies should explore a higher number of users of varying physical and cognitive ability to dive deeper into these immersive effects between the iVR mediums. Additionally, more systems should be explored beyond the Mechdyne CAVE Flex and the HTC Vive Pro 2018, especially with the deluge of mixed reality devices hitting the market such as Magic Leap, Microsoft Hololens, and more. While these are costly devices at the current moment, a similar trend with iVR technology may occur in the near future.

Furthermore, this study was not conducted in a clinical setting and did not utilize clinical-grade biometric sensors. Due to the physical constraints of the CAVE, users were tested onsite at UCSC with only caretakers present, although healthcare professionals were involved either through remote meetings, check-ins, or email correspondence rather than onsite at a therapy clinic. For our vision of the future, we hope to integrate these immersive experiences in each user's household, which will require this detachment from the clinical setting. With cost and user experience in mind, we chose to work with commercially available biometric sensors to collect biofeedback. This resulted in a lack of resolution from the biometric collection, where brainwave sampling was limited to the prefrontal cortex as opposed to clinical full head caps. In addition, heart rate was only collected through a single site optical sensor as opposed to clinical multi-site sticktrode locations. While these devices did not have the best resolution or sampling sites, the alternative would have been introducing higher setup times and discomfort for our users through costly sticktrodes, electrogels, and other materials required by these clinical-grade biometric sensors. It should be noted that other researchers are reporting success in using these commercially available sensors by implementing computational and sensor fusion techniques for better analysis [140, 121, 122, 123].

With these limitations in mind, we are preparing future experiments to address these challenges with various local healthcare organizations in Santa Cruz, California. The framework shown in this chapter for analyzing game performance, biometric response, and survey collection will be utilized in these upcoming studies to personalize and adapt the healthcare experience for users of varying abilities.

3.5 Conclusion

Modern iVR systems are becoming ever more prevalent in the consumer marketplace, thus it is critical to compare room-scale and HMD based iVR mediums. This study is one of the first to compare these iVR systems in the context of physical exercise. Our findings suggest that HMDs have finally caught up to and may have even surpassed CAVEs with our exercise game for both adults with and without disabilities. We also highlight a pipeline for multi-modal exercise analysis from game behavior, physical movement, and biometric response. These insights may be useful to future developers and engineers from system design, user experience, and data analytics perspective.

With a high number of VR systems commercially available and emergent immersive accessories being created, there are numerous platform options for experimentation by healthcare researchers. In the future, we hope to refine comparison measurements between iVR systems and address different populations of all abilities in iVR health applications. More studies must be conducted in comparing these systems, especially with the goal of addressing a greater variety of healthcare issues. One possible future application for healthcare is where users connect virtually with therapists for evaluation, perform gamified task-based objectives to meet exercise goals, and use analytics to adjust the difficulty and speed of prescribed exercises.

Over fifty years ago, Ivan Sutherland demonstrated the first iVR HMD to the world [141]. For Ivan Sutherland, his vision of the future of iVR was one of an ultimate display: "the ultimate display would, of course, be a room within which the computer can control the existence of matter. A chair displayed in such a room would be good enough to sit in. Handcuffs displayed in such a room would be confining, and a bullet displayed in such a room would be fatal. With appropriate programming, such a display could literally be the Wonderland into which Alice walked" [142]. Modern iVR systems are enabling deeper and more rich experiences of presence into the virtual world [21]. These elements of the ultimate display that influenced modern iVR as we know it appears to be near at hand [143]. Given this progress of immersive mediums into the virtual world, we ask what would the ultimate iVR system be for exergaming and health?

The study presented in this chapter supports that a modern HMD such as the HTC Vive is more engaging and produces better physical exercise performance than the more expensive room-scale CAVE medium. Through Project Star Catcher and its framework, we have explored the effects of the virtual world for individuals with and without disabilities [23]. Through our comparative study, we have seen that modern HMDs have a vast potential for physical exercise games for users of mixed abilities. In addition, through integrating biofeedback and motion capture analytics, iVR healthcare experiences can be personalized to match the needs and motivations of the user. With growing advances in artificial intelligence and machine learning, perhaps future iVR exergames can learn from both the users and therapists to best prescribe and augment VR stimuli for exercise. We envision this medium to be one that adapts the virtual world to the run-time emotional and physical state of each user to create a profound and maximally engaging experience. This chapter of research was published in the journal *ACM Transactions on Computing for Healthcare* and titled "On Shooting Stars: Comparing CAVE and HMD Immersive Virtual Reality Exergaming for Adults with Mixed Ability" [144].

Chapter 4

Virtual Reality for Physical Therapy

4.1 Introduction

Our previous work showed that we should design a rehabilitation game for a head-mounted display system considering user performance and physiological response. Our goal for this game was to make exercise more entertaining and engaging since many patients stop their at-home exercises once the accountability of in-person therapy sessions ceases. This means these patients never fully recover as rehabilitation is seen as a continuum of recovery and maintenance by physical therapists. We designed and created the game while collaborating with physical therapists to ensure patient safety. This chapter examines an iVR-based experience for upper-extremity rehabilitation called "OpenButterfly," where users follow movements to protect a virtual butterfly. OpenButterfly enables a dynamically controllable environment for individual exercise by utilizing motion capture, a biomechanical model of torque and angular momentum, and a biometric pipeline for brainwave, heart-rate, and skin conductance analysis.

The goal of this study was to evaluate our iVR therapy system over the course of 8 weeks. Our target users are those recovering from musculo-skeletal injuries who have completed conventional therapy and need to continue therapy exercises without the monitoring of a therapist. We performed our study in a lab setting to evaluate user performance and experience through musculoskeletal simulation, game analytics, questionnaires, and biofeedback response. We coined this new system of rehabilitation as OpenButterfly. OpenButterfly is a heavily modified version of Project Butterfly by Elor et al. [136]. We also developed a new game tool to easily record and implement custom exercise movements into the game, run repetitive personalized exercise sets with individual users, and have developed a pipeline for multi-modal analysis. Specifically, the goals of this study were to evaluate the following:

1. User performance using game play analytics.
2. Forces and total movement during gameplay sessions using biomechanical simulation.
3. User experience by measuring physiological response to gameplay and gather user feedback.

Through this work, we hoped to highlight methodologies for other researchers interested in diving deeper into the rehabilitation process with immersive virtual environments, biomechanical analysis, and biofeedback.

4.2 Experimental Design

Our target user group consisted of outpatients recovering from shoulder injuries that were pre-cleared for participation. Additionally, these users were patients who failed to complete their at-home exercises, which, as explained prior, can lead to incomplete recovery and increase the risk for re-injury. To recruit study participants, a survey was emailed to interested university students with general questions about their desire to participate in the study, if they had a relevant injury, if they participated in physical therapy, and what stage of recovery they were currently in. Follow up interviews were conducted with respondents to get more information about their injury and long term recovery goals to determine if they met the user group criteria. After such screening, five students (one female, four males) with ages ranging from 21 to 28 were chosen, and each student provided informed written consent to participate in both studies. All participants were continuing normal daily living activities but claimed to have limited strength, and/or a limited range of motion. This study received IRB approval from the Office of Research Compliance at the University of California - Santa Cruz. To document user participation, we established a data-collection pipeline and methodology.

4.2.1 Methodology and Data Collection

The effects of OpenButterfly were examined during an eight week period through a multimodal analysis of biomechanical, biometric, and gameplay responses. To enable such an analysis, we designed the OpenButterfly pipeline, as shown in Figure

4.1, and applied the analysis after testing individual users per exercise session. Our user testing protocol can be seen in Figure 4.1. The sessions consisted of the following methodology stages:

1. Preparation: The study administrators sanitized the iVR equipment, made sure all biometric sensors were fully charged and ran a session of OpenButterfly with all sensors active to ensure data communication quality.

2. Baseline: All biometric sensors were placed on the user in the exercise area. The administrator instructed the user to remain still and relax. After a 15 second period of adjustment, a 30-second baseline was recorded to mark each users' resting state for every session.

3. Rest: The user was instructed to relax for 90 seconds before performing the exercise with OpenButterfly. This was done before every new exercise was prescribed.

4. Exercise: Users completed 60 seconds of gameplay using OpenButterfly with the iVR headset and biofeedback/biomechanical data recording system. Upon completion of one set, the Rest stage was repeated.

5. Survey: After all exercises were completed, users filled out a brief survey indicating preference, pain, immersion, and self-reported advancement toward long term recovery goals. Such survey questions can be seen in Table 4.3.

Two researchers were always present to monitor user experience and followed a strict written protocol when interacting with users. This ensures a consistent method of tracking progression in the course of the eight weeks. Below is the list of equipment that we used in the study. Every sensor was chosen with accessibility and cost as a factor.

1. HTC Vive: Through Vive and the Unity Game Engine, motion capture and game data are recorded during runtime at 90 Hz using a data exportation method developed in previous studies by Elor et al. [136, 23].

2. OptiTrack: A motion tracking system of 10 reflective markers was recorded at 120 Hz using 8 Optitrack 13W cameras [145].

3. InteraXon Muse 2 - Brain Sensing Headband [119]: Muse is a commercially available headband that records EEG on the pre-frontal cortex (TP9-10, AF1-7) with dry contact electrodes. The headset has built-in internal noise suppression with 2uV RMS noise floor and a 50 or 60Hz regional notch filter. Muse was connected wirelessly to the Muse Monitor app on a mobile device and employed a Cooley-Tukey FFT [146] to extract brainwave band power in bels. Muse has successfully been used in other studies to infer mental state, analyze event potentials, and record biofeedback [122, 123, 121]. Foreheads were sterilized with saline wipes before gameplay.

4. Neulog GSR logger sensor NUL-217 [118]: The NUL-217 is a GSR logging sensor that measures the conductivity of the skin between the fingers. The logger connects to a USB-200 Module and records GSR in micro Siemens with a 10nS resolution at a max sample rate of 100Hz. The two finger electrodes were sterilized

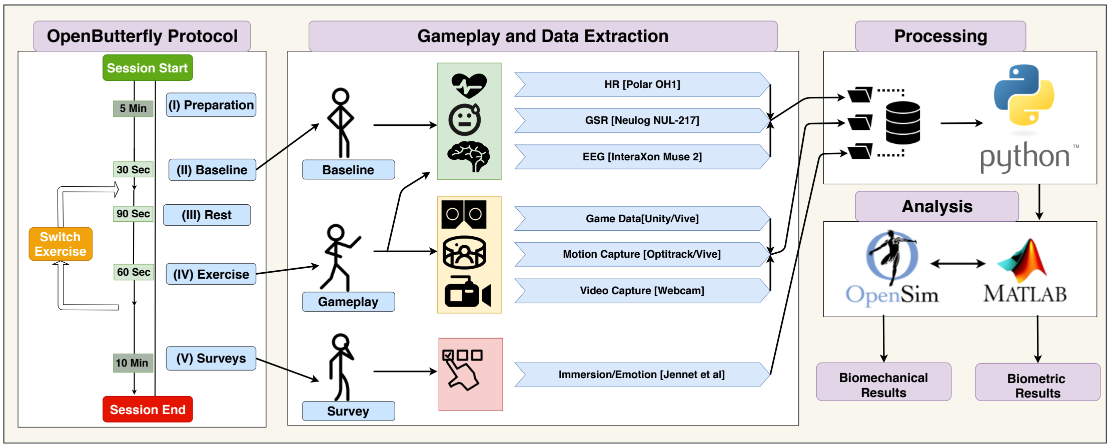

Figure 4.1: OpenButterfly Protocol & Data Pipeline Illustration for both Pilot [A] and Revised [B] Studies. OpenButterfly Protocol indicates the general outline for each experimental session. As shown in Gameplay and Data Extraction, the HR, GSR, and EEG were independently collected for a baseline, and then collected with game data, motion capture, and video capture during gameplay. Our survey was administered at the end of each session. After each session, we compiled all the data files through synchronization achieved via Python. MATLAB R2018B [1] was used to run statistical analysis on biometric data, and OpenSim [2] utilized the tracking data to calculate shoulder joint kinematics and dynamics.

with saline wipes before gameplay.

5. Polar OH1 - optical heart rate sensor [117]: the Polar OH1 is a 6 LED optical heart rate sensor that is used with an armband to record beats per minute through Bluetooth at 1Hz sampling frequency.

These devices are easy to set up for a user at home and are a more affordable solution compared to clinical grade sensors; i.e. these biofeedback sensors do not utilize single-use components, unlike more conventional systems that may use EEG gels or sticktrodes.

4.2.2 Game Mechanics

Our game, titled “OpenButterfly”, consists of a virtual butterfly that moves within reach of the participant to guide the user through their required movements. OpenButterfly gameplay can be seen in Figure 4.2. It is an adaption of Project Butterfly by Elor et al, previously designed for upper-extremity impairments for older adults while using a soft robotic wearable [136]. Specific new contributions to the OpenButterfly software includes a new system that records and prescribes custom exercise movements, runs automated repetitive personalized exercise sets with individual users, and provides increased stimuli for movement guidance. These contributions were designed through feedback sessions with collaborating physical therapists across Santa Cruz, California.

To guide movements, projectile crystals emanate from a 15m distance and move on a collision path with the butterfly. Users were informed that the goal of the game is to protect the butterfly from these crystals. The player holds an orb in their hand that they can place over the butterfly to protect it. The crystals explode when they hit the orb, letting the player know they successfully protected the butterfly and earned a point.

For prescribing custom movements, the path of the butterfly can be predetermined and set using a simple interface. The therapist can enter the ”Path Development” game mode, where they see the butterfly in their hand. When the trigger is pressed, they can move the butterfly in any path they desire at any speed. These movements can be saved and accessed through internal comma-separated value files. Through these movement files, the butterfly will follow recorded exercises that are automatically normalized to each user's arm length, target arm, and prescribed speed of movement. These implementations were done through utilizing the Unity3D Game Engine's Microsoft .NET File I/O C# Libraries [147].

Our study examined this new game mode by recording and prescribing seven new exercises in collaboration with therapists. The distance of the butterfly from the user is scaled based on arm length, which was measured for each participant at the beginning of the study using the relative position of the game controllers to the headset. Thus the game automatically scales exercise paths to the user's arm length – a feature requested by our collaborating physical therapists. Such game paths can be seen in Figure 4.3. These changes were done as a means for therapists to adjust the game to fit their users' needs and to enable dynamic customization and calibration. While some of the game's stimuli have been explored by Elor et al. through Project Butterfly, OpenButterfly was one of the first studies to examine these new exercise features when applied to iVR therapy over the course of eight weeks.

4.2.3 User Feedback

At the end of each session, participants were asked several Likert scale questions about their iVR experience that day. These questions were taken from a Jennett et al. survey for immersion in games [126] and was modified to focus more on user engagement. Such survey questions can be seen in Table 4.3. The surveys were utilized to track self-perception for users at the end of each session. These questions were used

Figure 4.2: A view of OpenButterfly gameplay. The protective transparent blue orb is outlined in white. The purple arrow shows the next incoming crystal cluster that heads towards the butterfly. To earn points, users need to place the orb over the butterfly to protect it from the crystal. Each crystal that is blocked earns a point.

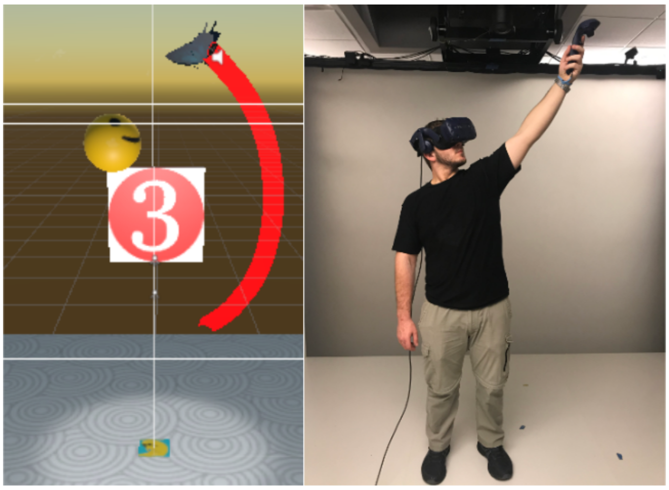

Figure 4.3: The "Path Development" custom game mode for therapist movement implementation. The right picture showcases a researcher using the iVR control to trace the path of the butterfly. The left picture indicates the movement's path, traced in red, so the researcher can see where exactly the path is located.

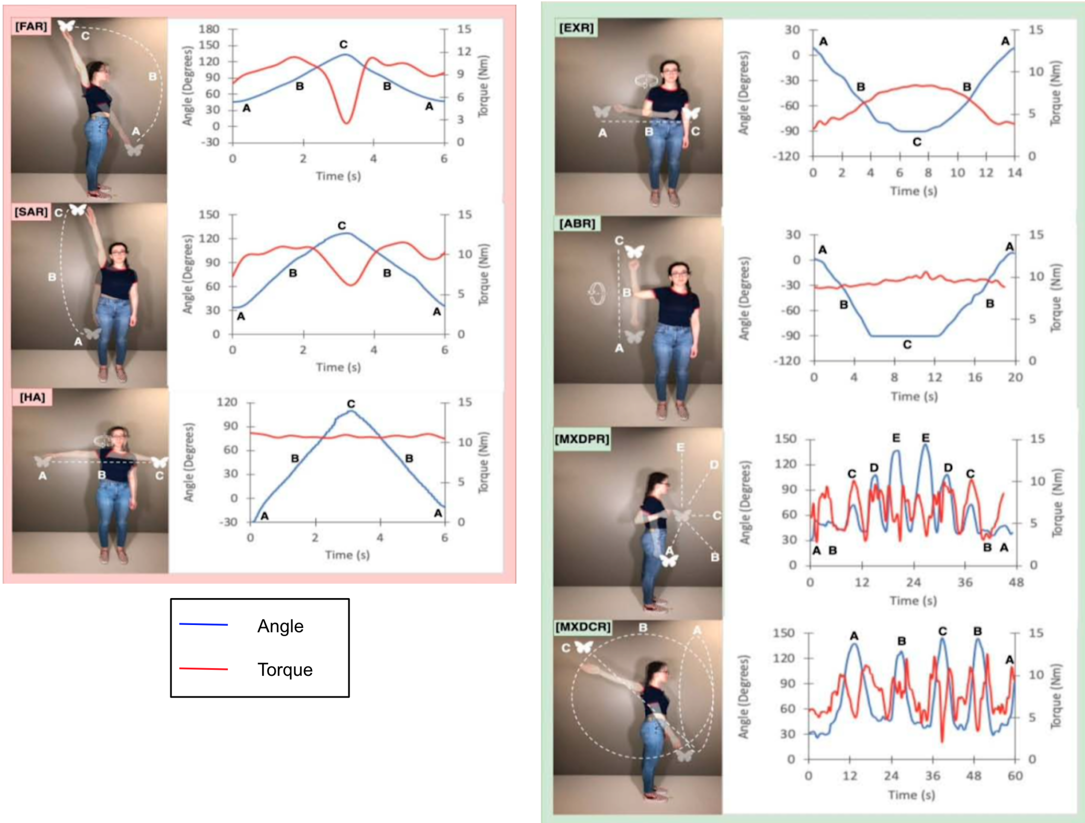

Figure 4.4: OpenButterfly Movements and OpenSim Outputs are shown above. The pilot study included the movements with the red background, while the revised study included both the red and green background movements. The movements are: FAR = Forward Arm Raise, SAR = Side Arm Raise, HA = Horizontal Abduction, EXR = External Rotation, ABR = Abducted Rotation, MXDPR = Mixed Press, and MXDCR = Mixed Circles. The white dotted line shows the path the butterfly traveled for each movement. On the graphs, the blue line shows the relevant angular displacement of the shoulder, and the red shows the torque placed on the shoulder throughout the movement.

to evaluate if the users would remain engaged, entertained, and immersed over the eight week period. Additionally, an exit interview was conducted at the end of the eight weeks to determine what modifications users would want to help improve rehabilitation. This enabled us to establish a mixed-method approach of gameplay, biomechanical, biometric, and survey responses for OpenButterfly.

4.2.4 Data Processing

Each of the biosensors, the Optitrack motion capture system, and HTC Vive produced their own output files with their respective recording frequency. Approximately 1,200 data files were produced during the Pilot and Revised Study. Python [148] scripts were written to structure the file management system and then sync all the files for each user session. OpenSim [2] simulations were ran using the motion tracking files from the Optitrack system, generating approximately another 1,100 files. Statistical analysis was then conducted on the data files using MATLAB [1]. The full pipeline for collecting and processing our data can be seen in Figure 4.1.

4.3 Pilot Study

The goal of the pilot study was for participants to perform common daily movements with an incremental and gradual amount of weight increase on their arms. The movements chosen were Forward Arm Raise (FAR), Side Arm Raise (SAR), and Horizontal Abduction (HA). These movements are simple single plane movements. We were careful to start with low-intensity movements to ensure participant's safety. Additionally, we hypothesized that participants would initially be unable to have full ROM through each exercise, but would be able to progress to full ROM with a small amount of wrist weight by the end of the pilot study.

4.3.1 Protocol of the Pilot Study

The pilot study was performed for the first four weeks. Each week consisted of two sessions where users performed 30-45 minutes of exercise (time includes rest). During each session, exercises were performed in the following order: FAR, SAR, and HA in order for a total of three rounds. The first round was played without weight for a warmup, and the subsequent two rounds were performed with the appropriate weight per user. Additionally, users had a 90-second rest between exercises, and each exercise was performed for 60 seconds at ten repetitions per minute. Aspects of this protocol are highlighted in Figure 4.1.

Full ROM for these movements was attained before adding weight to the user's wrist. The weight was added in small increments to elicit a strength progression, and users' average weight per session can be seen in Table 4.1. Weight was only increased when participants could comfortably perform two consecutive rounds of all three exercises for a given weight. To account for the participant's responses being influenced by the novelty of the VR game and or headset, an initial session was performed to introduce the game mechanics and enable the participants to be familiar with the OpenButterfly environment and movements.

4.3.2 Results of the Pilot Study

The averages of collected data can be seen in Table 4.1. The most prominent observation was that all users were able to complete the entire ROM of each exercise quickly, indicated by the high compliance rate, which allowed us to begin using weight early on in the study. To understand user engagement, effort, and immersion we employed a modified survey from Jennet et al. [126]. Table 4.3 shows the questionnaire asked at the end of each user testing session. Generally, users agreed that the game was engaging; they put a lot of effort into participating and felt immersed throughout gameplay sessions.

All elements of player behavior and biometric events (with the exception of user jaw clenches) were found to be significant, as seen in Table 4.1 A-s1 to A-s5. Users were able to acclimate to a 100% increase in weight resistance while moving their weak arm at a total average of 30m of distance per session. Additionally, users were able to successfully protect the butterfly at a compliance rate of a mean 96%, where compliance is defined as the time protecting the butterfly divided by the total time of the exercise session. In terms of both compliance and arm travel distance, these results held a low range of standard deviation, indicating that user performance was fairly constant between all users for these sessions. From the biometric data starting at a resting heart rate, the exercise sessions induced an average mean increase of 11.78 beats per minute, indicating increased physiological intensity from the exercises (shown in Table 4.2 [A]). Galvanic skin response was also found to be at a positive increase for all pilot protocol sessions, with a mean 1.47 micro Siemens resting-state change indicating arousal from gameplay, as shown in Table 4.2. For brainwave activity, the pilot protocol generally held a mean increase of all wavebands for alpha, beta, delta, theta, and gamma powers – this may confirm that users were mentally stimulated and physiologically challenged during gameplay.

4.3.3 Influence on the Revised Study

One thing that we learned from the pilot study is that the exercises of the Pilot Study were effective in increasing general strength, as can be seen in Table 4.1 where average weight between each session increases consistently. However, our users had a more substantial initial ROM than we anticipated. For our revised study, we needed to help our users progress more in ROM than the exercises in the Pilot Study required.

We performed ROM expansion by adding two common adduction/abduction movements: External Rotation (EXR) and Abducted Rotation (ABR), as well as two multiplanar movements: Mixed Press (MXDPR) and Mixed Circles (MXDCR). These movements can be seen in Figure 4.4 in the green region. Since these new movements focused more on stretching, no weight was used while performing these four movements.

In the Pilot Study, the first round was always played without weight as a warmup. Since the new stretching games were played without weight, we decided to do two rounds first of the non-weighted movements followed by two rounds of the weighted movements. This is further explained in the Revised Protocol Section.

4.4 Revised Study

Learning from the pilot study results, we modified our game to have a more appropriate protocol for our users. To address insufficient ROM exercises, new exercises were created, as shown in Figure 4.4 (EXR, ABR, MXDPR, MXDCR). These movements require a greater ROM at different angles than the pilot games. FAR, SAR, and HA games were kept to specifically address increases in strength by still utilizing the wrist weight progression protocol.

4.4.1 Revised Protocol

The revised study lasted four weeks, with twice a week sessions consisting of 30-45 minutes of exercise. During each session, two rounds of EXR, ABR, MXDPR, and MXDCR were performed with a one-minute rest between each exercise. Each movement was performed for 60 seconds at a slow tempo to allow for stretching at the limit of each subject's ROM. These exercises were always performed without any wrist weight as stretching was the goal, not strength.

Afterward, two rounds of FAR, SAR, and HA were performed with a one-minute rest between each exercise. This followed the same weight increase protocol as the pilot study to ensure a safe progression in strength exercises.

4.4.2 General Results of the Revised Study

From the revised protocol, users engaged in greater weight resistance than the Pilot Study [A], and subsequently, there was far more variability between users. Arm travel distance was less in total, but the movements were far more complex and slower. HR and GSR were found to be less than the Pilot Study's [A] mean resting state change but still elevated by nearly 10bpm and 1.13 micro Siemens, respectively. The lowered heart rate may be an artifact of the slower tempo in movement, and declining GSR may further indicate acclimation to the game with a lowering rate of arousal (however, it was still elevated far above resting state). Table 4.2 lists these results. As this table shows, the Revised Study [B] results were significantly different from the Pilot Study's results [A] for all data sets. Specifically, the Pilot Results [A] had a greater compliance rate, weak arm movement, HR change, and GSR change. In contrast, the revised study's results show higher levels of brain activity for all wavebands as well as Blinks and Jaw Clenches. This may indicate that the Revised Study was more mentally stimulating while both increasing weight resistance and game compliance, as shown in Table 4.1.

4.4.3 Biomechanical Performance

Using OpenSim with the motion tracking data from each session, we were able to determine the amount of torque placed on the shoulder for each exercise, as shown in Figure 4.5. We took the integration of torque with respect to time to determine the amount of angular momentum the shoulder generated from exercise.

While the average torque for each user dropped between the Pilot Study [A] and the Revised Study [B], the average angular momentum for each user increased between the studies, as shown in Table 4.2. This decrease in mean torques is a result of more exercises being performed without weight. For example, 8 of the 14 exercises in

Variable Averages	A-s1	A-s2	A-s3	A-s4	A-s5
Weight Resistance [lbs]	1.80 (0.73)	2.05 (0.88)	2.69 (1.16)	2.69 (1.65)	3.19 (1.50)
Torque [Nm]	10.00 (0.23)	9.87 (0.28)	10.02 (0.20)	10.17 (0.44)	10.24 (0.56)
Angular Momentum [kNms]	5.40 (0.13)	5.32 (0.15)	5.41 (0.11)	5.49 (0.24)	5.49 (0.36)
Compliance Rate [%]	96.45 (1.51)	95.08 (2.74)	95.31 (1.18)	96.99 (1.05)	95.12 (1.51)
Arm Traveled [m]	30.46 (0.48)	30.24 (0.32)	29.67 (0.60)	29.41 (0.53)	29.56 (0.50)
HR! [bpm]	12.05 (4.27)	9.49 (2.51)	11.81 (3.58)	14.03 (4.59)	11.53 (2.78)
GSR! [uS]	1.62 (0.75)	1.28 (0.68)	1.64 (0.98)	1.54 (0.70)	1.32 (0.56)
Alpha Power! [bels]	0.05 (0.04)	0.11 (0.05)	0.13 (0.04)	0.12 (0.03)	0.10 (0.03)
Beta Power! [bels]	0.18 (0.03)	0.20 (0.08)	0.19 (0.04)	0.29 (0.06)	0.17 (0.04)
Delta Power! [bels]	0.12 (0.07)	0.36 (0.15)	0.35 (0.08)	0.38 (0.17)	0.19 (0.08)
Theta Power! [bels]	-0.04 (0.04)	0.17 (0.05)	0.16 (0.05)	0.24 (0.13)	0.08 (0.05)
Gamma Power! [bels]	0.30 (0.06)	0.23 (0.10)	0.29 (0.05)	0.47 (0.05)	0.29 (0.05)
Blinks! [per s]	-0.48 (0.06)	0.12 (0.10)	0.02 (0.03)	-0.19 (0.12)	-0.16 (0.11)
Jaw Clenches! [per s]	0.01 (0.07)	0.01 (0.02)	≈ 0 (≈0)	≈0 (≈0)	0.01 (0.04)
Variable Averages	**B-s1**	**B-s2**	**B-s3**	**B-s4**	**B-s5**
Weight Resistance [lbs]	2.94 (1.62)	3.55 (1.73)	3.80 (1.69)	4.19 (1.71)	4.55 (1.93)
Torque [Nm]	8.55 (0.21)	8.43 (0.31)	8.63 (0.62)	8.45 (0.36)	8.57 (0.69)
Angular Momentum [kNms]	7.14 (0.22)	7.01 (0.27)	7.33 (0.49)	7.10 (0.30)	7.21 (0.55)
Compliance Rate [%]	90.14 (6.26)	91.06 (6.44)	94.43 (3.95)	94.43 (3.92)	95.58 (3.48)
Arm Traveled [m]	21.14 (9.041)	20.93 (8.74)	20.35 (8.65)	20.44 (8.64)	20.58 (8.97)
HR! [bpm]	7.55 (4.20)	7.59 (3.09)	8.67 (5.00)	12.82 (6.83)	12.65 (6.59)
GSR! [uS]	1.37 (0.75)	1.29 (0.93)	0.99 (0.92)	0.93 (0.68)	1.07 (0.77)
Alpha Power! [bels]	0.19 (0.12)	0.15 (0.18)	0.06 (0.11)	0.08 (0.07)	0.10 (0.09)
Beta Power! [bels]	0.34 (0.14)	0.26 (0.22)	0.25 (0.13)	0.19 (0.15)	0.13 (0.14)
Delta Power! [bels]	0.44 (0.25)	0.36 (0.38)	0.36 (0.23)	0.19 (0.16)	0.28 (0.22)
Theta Power! [bels]	0.24 (0.18)	0.24 (0.18)	0.15 (0.14)	0.09 (0.09)	0.10 (0.11)
Gamma Power! [bels]	0.50 (0.23)	0.42 (0.32)	0.42 (0.19)	0.34 (0.22)	0.23 (0.19)
Blinks! [per s]	-0.09 (0.30)	0.24 (0.40)	0.15 (0.17)	0.07 (0.24)	-0.11 (0.32)
Jaw Clenches! [per s]	0.02 (0.02)	0.04 (0.05)	≈0 (0.03)	≈0 (≈0)	-0.01 (0.04)

Table 4.1: OpenButterfly Protocol Pilot Study [A] and Revised Study [B] session averages between all exercises for all users. Parenthesis indicated standard deviation. Exclamation Mark indicates resting-state change (note all biometric measurements indicate change induced from gameplay compared to baseline measurements).

Variables [Na=225 & Nb=350]	Sig	[A] Pilot Mean (STD)	[B] Revised Mean (STD)
Weight [lbs]	***	2.55 (0.505)	**3.90 (0.495)**
Torque [lbs]	***	**10.06 (0.15)**	8.53 (0.08)
Ang. Momentum [kNms]	***	5.42 (0.07)	**7.18 (0.10)**
Compliance [%]	***	95.26 (0.244)	**95.61 (1.418)**
Arm Traveled [m]	***	**29.87 (0.103)**	20.69 (0.182)
HR[a] [bpm]	***	**11.78 (0.906)**	9.86 (1.584)
GSR[a] [uS]	***	**1.47 (0.155)**	1.13 (0.110)
Alpha[a] [bels]	***	0.10 (0.008)	**0.12 (0.041)**
Beta[a] [bels]	***	0.21 (0.020)	**0.23 (0.036)**
Delta[a] [bels]	***	0.28 (0.049)	**0.33 (0.085)**
Theta[a] [bels]	***	0.12 (0.035)	**0.16 (0.040)**
Gamma[a] [bels]	***	0.32 (0.021)	**0.38 (0.054)**
Blink[a] [per s]	**	-0.14 (0.038)	**0.05 (0.086)**
Jaw[a] [per s]	***	0.005 (0.028)	**0.011 (0.018)**

Table 4.2: Wilcoxon Significance For Pilot Study [A] Vs. Revised Study [B] Results. The Protocols Were Found To Be Significantly Different From Each Other At 95% Confidence In All Data Categories. "Sig" Indicates The Significance Level. Superscript (A) Indicates Resting-State Change (Note All Biometric Measurements Indicate Change Induced From Gameplay Compared To Baseline Measurements). Bolded Values Indicate Significant At 95% Confidence From Wilcoxon Testing. Pilot Study Indicates Higher Game Performance As Well As GSR And EEG.Revised Study Shows Higher EEG Performance As Well As Blinks And Jaw Clenches. Note That Na And Nb Is The Total Number Of Samples Found By Number Of Sessions × Number Of Users × Number Of Exercises

the Revised Study [B] were played without weight, whereas only 3 out of the 9 exercises in the Pilot Study [A] were without weight. The increase in average angular momentum between the Pilot Study [A] and the Revised Study [B] occurs because more games are played in the Revised Study [B]. Figure 4.5 shows each user's average torque and angular momentum for every session.

We expected to see more of a steady increase for torque and angular momentum in each study since the users were lifting the same or more weight than their previous session. However, we observed that users' average torque and angular momentum fluctuated a bit from session to session in the same protocol rather than continuously increasing. We believe this is important to show as we can see which sessions the users were not performing as expected. Users told us that some days they would come in more stiff or sore than other days. This data may be indicative of these cases and theoretically could be used to help adjust the exercise protocol. These results may be helpful for other researchers interested in expanding upon this work for a more personalized and reactive therapy regime. Understanding day to day fluctuation through this data could lead to new algorithms for a more customizable rehabilitation through adapting to users' capabilities. Even if users decline or have a set back (i.e., sore from previous days activities such as yard work), these insights could be used to tailor the difficulty to the most important muscle groups for maximizing therapeutic benefit.

From this data, we can also see the most significant changes in torque and angular momentum come from the different amount of games played; in short, the more games played, the more angular momentum was gained. This is useful when deciding a

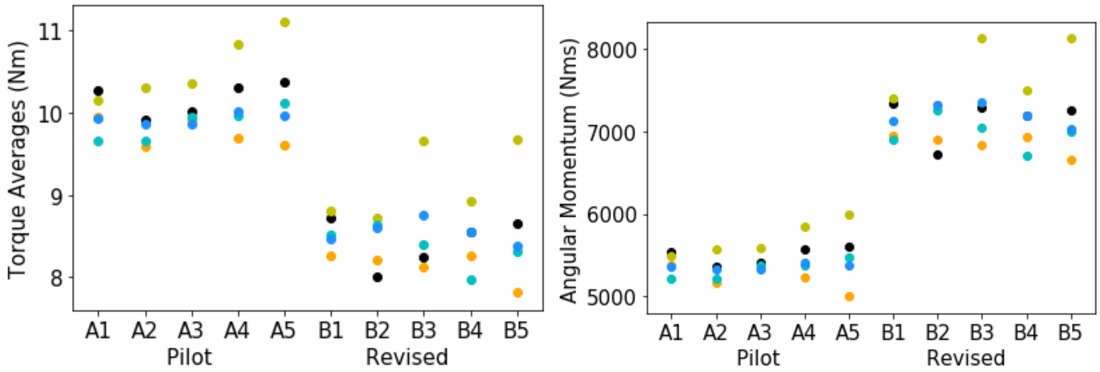

Figure 4.5: Average torque (top) and angular momentum (bottom) for each session is shown with each color representing an individual user.

rehabilitation plan as the variables to consider are weight resistance or total movement. So by utilizing the angular momentum, we can determine the exact amount of motion of a specific game, no matter the movement path of the arm. Such an analysis of angular momentum could be used to build a more intelligent progression.

4.4.4 Qualitative Performance

A significant difference between our two studies was found from Wilcoxon significance testing at 95% certainty for the qualitative user surveys, as shown in Table 4.3. Users reported the Revised Study [B] to be significantly more challenging ([Q7]), progressive ([Q10,Q11]), and liked ([Q1]) the protocol more than the Pilot Study [A]. Conversely, users reported that they felt the revised protocol provided significantly less distraction ([Q4]) from their real-life compared to the pilot protocol. These results may indicate that the protocol choices for the Revised Protocol [B] were successful in challenging each user from an engagement and effort perspective.

Post-Session Survey Questions	Sig	[A] Pilot Mean (STD)	[B] Revised Mean (STD)
[Q1] I liked playing the game	*	4.3 (0.79)	**4.6 (0.50)**
[Q2] The game distracted me from pain		3.7 (1.08)	3.9 (0.75)
[Q3] The game felt more engaging than my traditional therapy routine		4.4 (0.81)	4.4 (0.79)
[Q4] The game provides a distraction from my real life	**	**4.5 (0.79)**	4.0 (0.80)
[Q5] When I played the game, I felt I lived in the game world	*	3.2 (0.95)	**3.6 (1.13)**
[Q6[a]] I put a lot of effort into the game		4.4 (0.61)	4.4 (0.56)
[Q7[b]] How challenging did you find the game?	**	3.2 (0.74)	**3.8 (0.76)**
[Q8] I could still notice the outside world while playing the game		2.3 (0.83)	2.3 (0.76)
[Q9[a,d]] Did you ever want to quit playing?		4.8 (0.55)	4.7 (0.65)
[Q10[a]] Did you feel like you were making progress in the game?	*	2.9 (1.55)	**3.8 (0.88)**
[Q11[b]] How well do you think you performed in the game?	*	**4.1 (0.64)**	3.8 (0.62)
[Q12] Do you feel that you performed better than last time you played the game?		3.6 (1.03)	3.8 (0.79)
[Q13[c]] How much pain did you receive (feel) while you played the game?		2.6 (1.48)	2.2 (1.94)
[Q14[c]] How immersed did you feel when playing the game?		7.7 (2.55)	7.5 (2.45)

Table 4.3: OpenButterfly Survey Table. Results Without Asterisks Are In Likert Type Scale Where One Indicates Strongly Disagree And 5 Indicates Strongly Agree. "Sig" Indicates Wilcoxon Significance Level. Superscripts Indicate: (A) Scale Of "Not At All" To "A Lot", (B) Scale Of "Very Poor" To "Very Well", (C) Ten-Point Likert Scale For "Not At All" To "A Lot", (D) Indicates It Was A Reverse Question And The Response Average Is Represented In The Inverse To Keep All Values On The Same Scale.

4.5 Discussion

From our multimodal analysis of our eight-week study, we show that OpenButterfly accomplishes our goals of increasing ROM and increasing strength. This was a multi-step process that required two stages to adjust and further customize the game to the users' capabilities as rehabilitation necessitates. The Pilot Study [A] was useful to help determine the capabilities of our users and how to set achievable goals for them. This stage showed that the users achieved a full ROM for the first three exercises (FAR, SAR, HA) and were ready to start training with weights very quickly. Starting with simple movements was a safeguard against exercises that were too advanced for their state of recovery. The insights from the Pilot Study enabled us to create more complex movements and continue to work on strength.

For the Revised Study [B], we created four new movements (EXR, ABR, MXDPR, and MXDCR) that were performed without weight to target enhanced ROM. At the same time, the original exercises from the Pilot Study [A] were carried over to focus on strength building. We saw ROM increased to meet these challenging and further-reaching movements indicated by the increase in the rate of the compliance recovery during the Revised Study [B], as shown in Table 4.1. Also, in Table 4.1, results from average weight indicated a successful increase in strength. The Revised Study [B] displayed an improvement where more exercise difficulty was leveraged to safely challenge the users. Users also enjoyed the new exercises and stated it was similar to "unlocking a new level in a game." This is a consideration as we move forward: making levels that are of different movements so that the game remains challenging and does not become repetitive.

Additionally, the users' physiological recordings and self-reported responses indicated that users were able to remain engaged with the game beyond the novelty effect period for the course of the eight weeks, as seen in Table 4.1 and Table 4.2. Consequently, technology like OpenButterfly may become a promising tool for addressing the problem of adherence to a rehabilitation program. The more enjoyable and engaging the program, the more likely users will continue the program. This adherence with OpenButterfly is particularly exciting, as other researchers may be able to utilize similar iVR physical therapy experiences for long-term treatment.

4.5.1 Contributions from OpenButterfly

Through our study, we believe several insights can be useful to game developers and researchers. First, we learned that the ability to easily and quickly create custom paths for arm movements during gameplay allowed us to efficiently adjust our games between the Pilot Study [A] and the Revised Study [B]. This adjustment needed to occur because all users were able to complete a full ROM with added weight within a few sessions. The Pilot Study [A] exercises proved valuable as a baseline ROM and for improving strength. The Revised Study [B] had more complex movements targeting ROM and kept the now proven original exercises for targeting strength. These game modifications were guided by our collaborating physical therapists to increase the difficulty of an appropriate progression in strength and ROM. The ability to record custom motion paths and normalize movements to each user's arm length and height proved to be a valuable tool.

Another useful tool was the biomechanical simulation, as it offered more in-depth analytics into user performance through analyzing performance during a session. In traditional PT, the therapist can monitor progression through measurements of ROM and strength, typically pounds lifted or level of a resistance band. Through our study with OpenSim, we are able to provide this data and, in the future hope to have everything streamlined so that no matter the movements performed, simple or complex, we can provide a thorough representation of the amount force placed on the working joint for a therapist to examine. With further user testing, perhaps researchers can build more sophisticated models for simulation that will provide individual muscle for training and rehabilitation for any of the movements performed in the game. This can help the therapist target specific muscles to aid in a focused recovery.

Additionally, biometric data from OpenButterfly may help researchers understand users' physiological responses to the iVR experience. This helps with recovery as more enjoyable user experience is likely to lead to better adoption of a rehabilitation program. Our data indicates users had higher brain activity for the Revised Session [B], which should be further explored in considering rehabilitation monitoring and game adaption for future studies. These metrics also provide a possibility for determining how much effort the user is putting into the game on a physiological level. Since this isn't a strenuous workout, we want to make sure users are working at an appropriate level. We learned that the levels we chose were enough to elicit a strength increase response, but not so much that user is at risk of injuring themselves.

We believe our study has shown this game's feasibility for helping with the recovery process, and fellow researchers, developers, engineers, and therapists may find aspects of our research useful for their endeavors. Collecting HR, GSR, and EEG may provide a deeper understanding of a user's engagement and physical effort with iVR exercises. This can help with game development in creating exciting experiences to help increase a user's desire to play the game. Biomechanical simulation can provide valuable metrics to a monitoring therapist and also give a progress log over an extended period of time. OpenButterfly itself shows that other games can be created to aid with recovery, and we suggest from guidance with our collaborating therapists that in future games, there is a way for a therapist to dictate the movements of the game easily, so it is customizable to the user's needs.

4.5.2 Study Limitations

There are several limitations that may impact generalizabilty of the results. The study examined five users with Openbuttefly, but in the future iterations we plan to recruit more users. While the study lasted eight weeks it would be helpful to understand more long term effect by conducting the study for 12-16 weeks since our goal is adherence of users. Additionally, we limited the frequency to two sessions per week to ensure adequate time for recovery, but having the users progress to three and four times per week could yield better benefits. While our long term goal is at-home use, we conducted our study in a lab to examine performance and user experience of our system. The next step would be apply what we learned from our system feasibility study and conduct an at-home user study.

4.6 Conclusions and Future Work

The purpose of OpenButterfly was to create an effective and feasible iVR physical therapy game to help users with shoulder injuries through multimodal rehabilitation analysis. This was accomplished by working with therapists and enabling game recordings to mimic movements found in physical therapy targeted at ROM and strength training. Through OpenButterfly, we present a novel study that is a long-term, customizable highly immersive virtual reality game for shoulder rehabilitation that analyzes physiological response and uses biomechanical simulation to identify the joint kinetics and dynamics. Working with therapists, we have identified useful tools and data sensors to aid in developing games targeted at recovery that we believe other serious game researchers will find helpful. This multimodal rehabilitation analysis will help with the next iteration of our game to ensure user engagement and that users are working at an appropriate threshold, not too intense to risk injury but difficult enough to elicit a physiological adaptation.

We explored two experimental studies: a Pilot Study consisting of three single everyday movements targeted at a basic ROM and strength, and a Revised Study that incorporated four new movements aimed at ROM from insights gained from the multimodal rehabilitation analysis of the Pilot Study. Our results indicate that users were able to overcome the novelty effect of iVR through extended exposure to gameplay over eight weeks. We were also able to measure heart rate, galvanic skin response, and electroencephalography while our users played the game, allowing us to understand their physical strain and emotional response while playing. With the motion capture data, we were able to determine the kinematics and dynamics of the shoulder during gameplay through biomechanical simulation. We believe this data would be useful for physical therapists as it helps quantify the forces of the joint for an entire session and would provide a method for remote therapists to quickly understand users' exercise session.

In the future, we plan to explore the design of new levels within the game that contain more complex and less predictable movements to challenge users physically and mentally. Our long term goal is to develop an at-home recovery game that is capable of providing meaningful game data remotely to the therapist. Subsequently, we plan to explore a more complex biomechanics model capable of identifying individual muscle force contribution to movements. The incorporation of runtime biomechanical models to identify muscle weaknesses may further aid in custom movements for an individual user to help maximize their recovery by ensuring the targeted muscles are being used for a given movement. We hope to deploy this system for at-home use to make OpenButterfly more accessible for users in need. This work was published in the journal *American Journal of Sports Science and Medicine* and titled " Openbutterfly: Multimodal Rehabilitation Analysis of Immersive Virtual Reality for Physical Therapy" [3].

Chapter 5

The Novelty-Effect of Immersive Virtual Reality

5.1 Introduction

The previous chapter's study also allowed for the examination of the Novelty Effect regarding iVR. The Novelty Effect is when users improve due to interest in the new technology (iVR in our case) rather than from the intervention itself. This is important to understand how users will engage with this platform at home over the long term. If users are only interested because iVR is new and entertaining then that feeling will wear off over time and adherence will fail once again.

The study reported in this chapter aims to answer the question: Can an iVR HMD experience maintain engagement beyond the novelty period and show continued rehabilitative improvement using multi-modal analysis when used as a physical therapy environment? To answer this question, we expand upon an iVR serious game for controlled physical exercise. The purpose of the updated design is to investigate improvements in physical performance using an iVR system by following protocols that are similar to conventional physical rehabilitation. Three outpatient physical therapists with doctorate degrees in physical therapy and over 40 years of combined professional experience, helped design the protocols used in this study to match exercises used in clinical settings. Through these three consultants, we learned that the principles of functional shoulder rehabilitation for late-phase recovery usually extend from 6 to 12 weeks of treatment to "(1) restore full range of motion and flexibility... and (2) increase strength, power, and endurance with exercises that stress core-based muscle synergy" [149, 150]. In this study, we extend these principles to stimulate a range of motion in the first four weeks and increase strength in the last four weeks.

Specifically, the contributions of this study are:

1. A demonstration that our iVR HMD based serious game system can be effective for physical rehabilitation.

2. An examination of methods towards maintaining engagement and motivation over extended period of time.

3. An assessment of the feasibility of using biometrics to complement the iVR game.

Please reference the previous chapter for system design and protocol as this chapter analyzes the same data set collected during the last chapter.

5.2 Results

Results from session data were post-processed using the Mathworks Matlab 2018b Statistics and Machine Learning Toolbox [129]. We examined user performance between every session for mean and standard error. In total, we collected 225 session exercises for the Foundation protocol and 350 session exercises for the Challenge protocol for every data type. Biometric signals were normalized from each user's baseline resting state to examine the changes induced by gameplay.

The game performance data shows general improvements in weight resistance over time with maintenance of compliance and exercise movement, as shown in Figure 5.1. On average, users were able to handle more weight resistance per exercise than the initial session, and while the compliance remains almost constant in the Foundation protocol, it increases significantly in the Challenge protocol. In both protocols, users were able to perform the same movements with a gradual weight increase.

For physiological performance, PBF was able to record and monitor elevated HR and GSR measurements when compared to resting-state for all sessions of each protocol. Figure 5.2 shows the changes from resting baselines and indicates that PBF always induced an elevated HR (indicating physical engagement) and stimulated GSR by 1uS or higher (indicating induced arousal). In the Foundation Protocol, users maintained a constant level of increased physical activity with a slow decline of arousal. In the Challenge Protocol, users had increased intensity of physiological activity with a considerable decline of arousal that eventually stabilized.

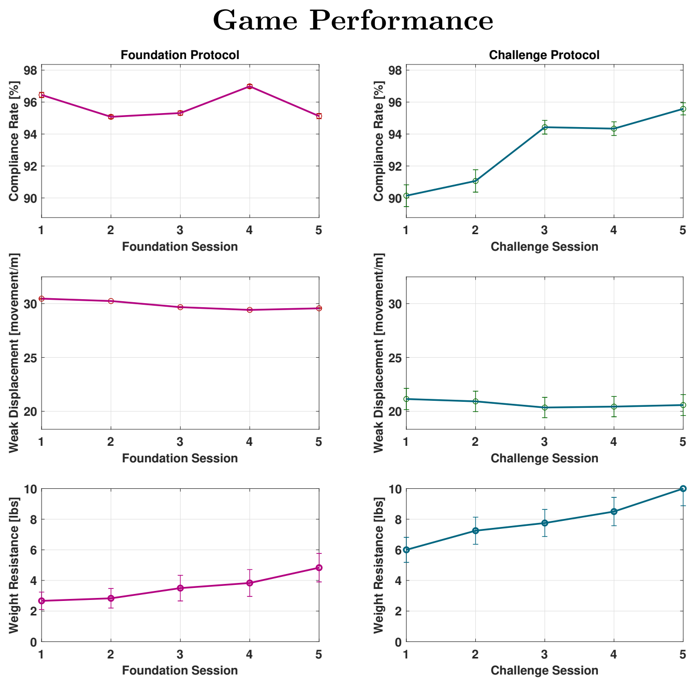

Figure 5.1: Game performance between Foundation Protocol (in red of 225 recorded exercises) and Challenge Protocol (in green of 350 recorded exercises). Row one shows compliance, where compliance is defined as the total time protecting the butterfly over the game's total time. Row two shows the mean upper-limb displacement between all exercises required in that session. Row three indicates the mean weight used between all exercises of that session. Error bars indicate standard error (note the Foundation Protocol had less variability between users, so error bars appear substantially smaller than Challenge Protocol due to shared scale).

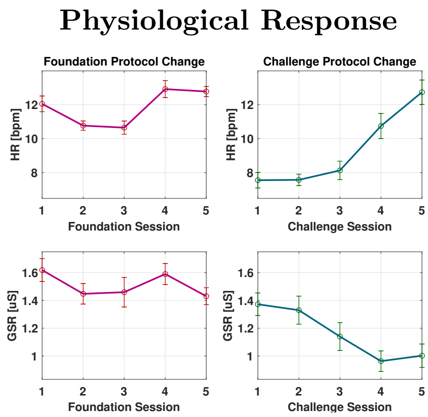

Figure 5.2: Physiological HR and GSR responses from gameplay are shown. Row one illustrates mean change from resting state of heart rate. Row two illustrates mean change from resting state of galvanic skin response. Biometric change is calculated as the offset between gameplay biometrics against resting-state biometrics. Error bars indicate standard error.

For brainwave response, neural activities were measures at all sessions and all protocols, as shown in Figure 5.3. All wavebands were found to be at a positive increase from resting-state change which indicates that Alpha, Beta, Delta, Theta, and Gamma waves were elevated during PBF usage. In the Foundation Protocol, all brainwave responses from users generally increased in the middle of the sessions and began declining towards the last sessions. In the Challenge Protocol, all brainwave activities had a more substantial initial session than the Foundation Protocol and generally declined overtime to nearly the same level as the Foundation Protocol's last session.

Additionally, Muse [119] holds the capability to detect facial muscle movements to determine a Boolean response of eye blinks and jaw clenches. This data was recorded during runtime gameplay, and converted to facial movements per second based on changes from the baseline, as seen in Figure 5.4. While playing PBF, users in the Foundation Protocol tended to blink less than their resting state for every session (with

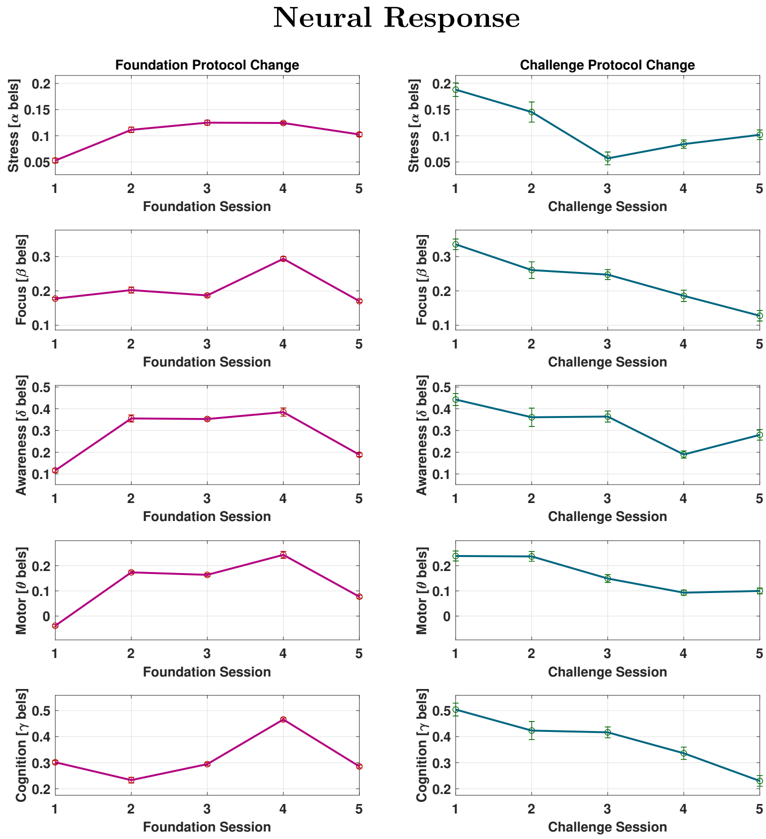

Figure 5.3: EEG responses between Foundation Protocol (in red of 225 recorded exercises) and Challenge Protocol (in green of 350 recorded exercises). Rows 1-5 show Alpha, Beta, Delta, Theta, and Gamma bands resting state change respectively. Error bars indicate standard error.

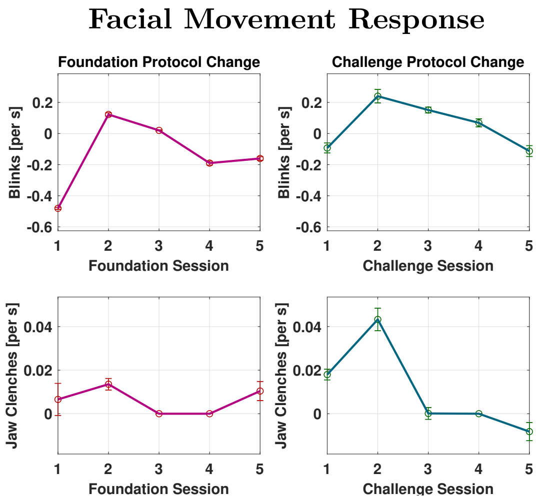

Figure 5.4: Facial muscle movements recorded with Muse between Foundation Protocol (in red of 225 recorded exercises) and Challenge Protocol (in green of 350 recorded exercises). Row one shows the mean resting state change of blinks per second. Row two shows the mean change of jaw clenches per second from resting state.

the exception of Session #3). Jaw clenches do not vary much between sessions. In the Challenge Protocol, users tended to blink and clench their jaw much more between every session than their baseline resting state. Unlike the Foundation Protocol, these blinks were always at a positive increase when compared to resting state, except for the first session, and were more rapid. Lastly, jaw clenches tended to decline as time progressed between sessions.

For user's self-reported responses, the qualitative survey questions can be seen in Figure 5.5 (engagement based) and Figure 5.6 (emotion based). For both protocols on each session, most users agreed that the game remained more engaging than their traditional therapy routine and that the game provided a distraction for them during their exercise, as shown in Figure 5.5. Similarly, the majority of users reported a positive range of emotions for each exercise ranging from Happy/Joyful, Excited/Motivated, and

Engagement Survey Response

Figure 5.5: Survey responses on engagement from 5 subjects, with 1=strongly disagree and 5=strongly agree.

Emotion Survey Responses

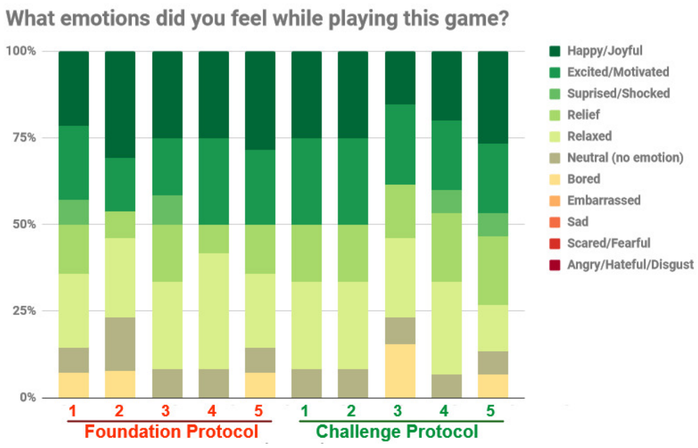

Figure 5.6: The self reported emotions ratios felt by users from post-gameplay survey.

Relaxed, as shown in Figure 5.6. Q3-4 show the largest differences in survey responses between protocols. Specifically on Q4, the Foundation Protocol saw a transition from unanimous disagreement with noticing "the outside world while playing the game" to a greater majority of neutral as time progressed. The Challenge Protocol was inverse to this effect, where users eventually became unanimous in disagreeing that they could notice the outside world during gameplay. In essence, this suggests that users were much more engaged in the game during the last two sessions of the Challenge protocol.

5.3 Discussion

Through analyzing the data from our two months study, we observed the following phenomena:

PBF was able to elicit rehabilitative responses similar to traditional therapy, including increases in muscle's strength, control and flexibility. The results suggest that across all users, their resistance successfully increased throughout the study, as evidenced by the weight increments that the users were able to cope with. Heart rate increased for both protocols, which we concluded were due to the increased weights that require additional muscular efforts. Compliance improved more during the Challenge Protocol than the Foundation Protocol, which may suggest that users that were challenged with the complex movements followed the protocol more carefully than asked to perform simpler movements. During the exit interview, users perceived that they gained significant strength and stability through playing the game. They felt they would have been unable to play the game at the beginning of the study using their final session's weights, and yet, users were able to perform those exercises with those weights.

Users can remain engaged in physical therapy using PBF and HTC Vive beyond the novelty effect period. One of the dangers of long-term therapy is when users get bored and lose interest in the exercises. We did not observe any decrease of interest and engagement beyond the novelty effect (when users were still new at iVR games). The greatest changes in all the brainwave bands (which are often associated with levels of stress [51], focus [52, 53], awareness [56, 57], motor [60, 61], and cognition [62, 67]) were seen in the transition from the Foundation Protocol to the Challenge Protocol. This sharp increase in all bands suggests that the additional exercises were able to engage the users considerably. Additionally, blinks were the lowest for each protocol's first session, indicating the user may have been more focused during these sessions [151]. This could mean that creating new types of movements after the user has become accustomed to a set of exercises can stimulate users to remain interested. When the Challenge Protocol was introduced, jaw clenches were increased from baseline, possibly indicating greater effort of the participant [152]. The survey responses also showed that users felt engaged by each protocol. In Q1 (Figure 5.5), users compared gameplay exercises to their traditional therapy. GSR responses declined over time in both protocols, and we speculate that this is most likely due to users becoming more acclimated to the game over time and thus causing drop in arousal. We should note that at the end of each protocol, GSR' level stabilizes, indicating a steady state arousal. The survey results suggests that PBF was more successful in enabling engagement of physical therapy than traditional interventions. Questions Q2-4 also demonstrated that the user felt present in the game world, possibly indicating a successful immersion. Additionally, users stated that they enjoyed the first few sessions, but started to lose enthusiasm as they felt the Foundation Protocol was too repetitive and straightforward. With the introduction of new and more complex movements that could not be easily memorized in Challenge Protocol, the users were excited once again to play the game. Some users stated that the more complex movements kept them engaged rather than "zoning out" as they did during the simple movements over time.

iVR games have the potential as a long-term physical therapy tool that can be used at home. PBF was successful in inducing rehabilitative response while maintaining extended engagement. The basic version of PBF only requires the Vive headset and none of the biosensors for remote usage. We argue that this suggests a low-cost solution for rehabilitation exercise compared to traditional long term physical therapy sessions that require users to visit a clinic.

It appears that the differences in the difficulty levels and goals between the two protocols induced noticeable change in the different brainwave measures. The most substantial changes were seen in the transition from Foundation Protocol to Challenge protocol, where new and more challenging games were added to the already existing games. From the context of Alpha band power (often associated with stress), these results may indicate that Challenge Protocol has much greater difficulty than Foundation Protocol, and induced a higher amount of stress when transitioning to new and more challenging exercises. It appears that Beta (often associated with focus) spiked higher when difficulty was increased, which was expected as users must focus harder when the game became more difficult. One take away message from these two findings is that, if we are to design games whose difficulty levels adapt to users' biometric changes, there should be a careful consideration to design how sharply difficulty levels increase, as increased levels of difficulty induces higher focus, but also higher stress.

More work needs to be done in establishing dynamic progression of difficulty, utilizing biomarkers, and testing more users. Progression during the game was an essential aspect for all users. They related adding new movements to the game to be like "unlocking a new level." All participants stated that they would like to have more levels to advance through and clear goals for each level. This would help keep things dynamic and avoid boredom due to repetition. Three of the five participants would recommend this game to a friend in its current version, while the last two stated they would recommend the game if more levels, progression, and goals were incorporated into the game.

It should be noted that future research should explore a more significant number of users and VR experiences to understand the long-term effects and user response of iVR physical rehabilitation gaming. As more immersive virtual environments are crafted for physical rehabilitation, there is a need to establish how such a system can be tuned to the user's biometrics to induce a desirable range of activity and understand how this will compare to conventional physical rehabilitation. In this study, seven different movements and one virtual environment were explored for upper extremity physical rehabilitation. More motions and varied experiences should be investigated to examine the game design, difficulty, and adaptation to iVR stimuli. We are also mindful that there were only five users that we followed for two months; however, we believe that this study is an important step towards gathering insights for future studies.

5.4 Conclusion

This study explored the effects of an immersive Virtual Reality HMD gamified upper-extremity physical therapy that record both physical and biometric responses over the course of two months. To provide a more engaging experience, we designed the study so that users completed their prescribed therapeutic movements by protecting a virtual butterfly in a dynamic and adaptive virtual environment. Two rehabilitative goals were set in the study: recovery of foundational movements and progressing with more complex motions. The study results suggest that movement improvements over time can be quantitatively assessed through game logs. The study also concluded that the biometric responses can complement game data and provide a richer insight on user engagement. These findings may indicate that long term immersive Virtual Reality physical rehabilitation is feasible.

In the future, we aim to expand PBF's capabilities for home health and to run larger trials for comparison with conventional therapy methods. We would like to dive deeper into the effects of immersive physio-rehabilitation through controlled trials to understand how user-perceived confidence and difficulty influences the recovery journey. Additionally, virtual environments, such as PBF, provide an opportunity to explore run-time biofeedback with adaptive difficulty using emotion classification, which we also plan to investigate. Motion capture data with biomechanical simulation may be utilized to estimate muscle forces for understanding biased movements and how to best prescribe rehabilitation towards addressing weaknesses. We plan to run biomechanical simulation for this estimation. The creation of an adaptive, personalized physical therapy game that adjusts to the user's mental and physical state in run-time may yield immense potential. The work examining the Novelty Effect was published in *IEEE Transactions on Games* and titled "Gaming Beyond the Novelty-Effect of Immersive Virtual Reality for Physical Rehabilitation" [153].

Chapter 6

Machine Learning for Physical Therapy

6.1 Introduction

With the development of Openbutterfly, we were awarded a National Science Foundation (NSF) grant and entry to the Innovation Corps Program (I-Corps). NSF aims to support innovation built upon fundamental research that can benefit society. I-Corps is an NSF program that provides researchers with mentoring and funding to accelerate innovation tailored to a group of end users who are in need of the newly developed technology. Through this program, we interviewed 130 physical therapists, our target end users, across the United States (Fig. 6.1) to learn what their biggest pain points were in their daily work. One of the most common issues brought up by physical therapists was the inability to easily get accurate biomechanical metrics (joint angles and forces) during remote telehealth visits. To address this problem, we developed machine learning models that can use the internal motion capture of an iVR system to output these types of metrics. We believe iVR coupled with our machine learning models address the issues brought up by physical therapists who were all using videoconferencing on laptops or phones. These models were built upon our previous work with Openbutterfly.

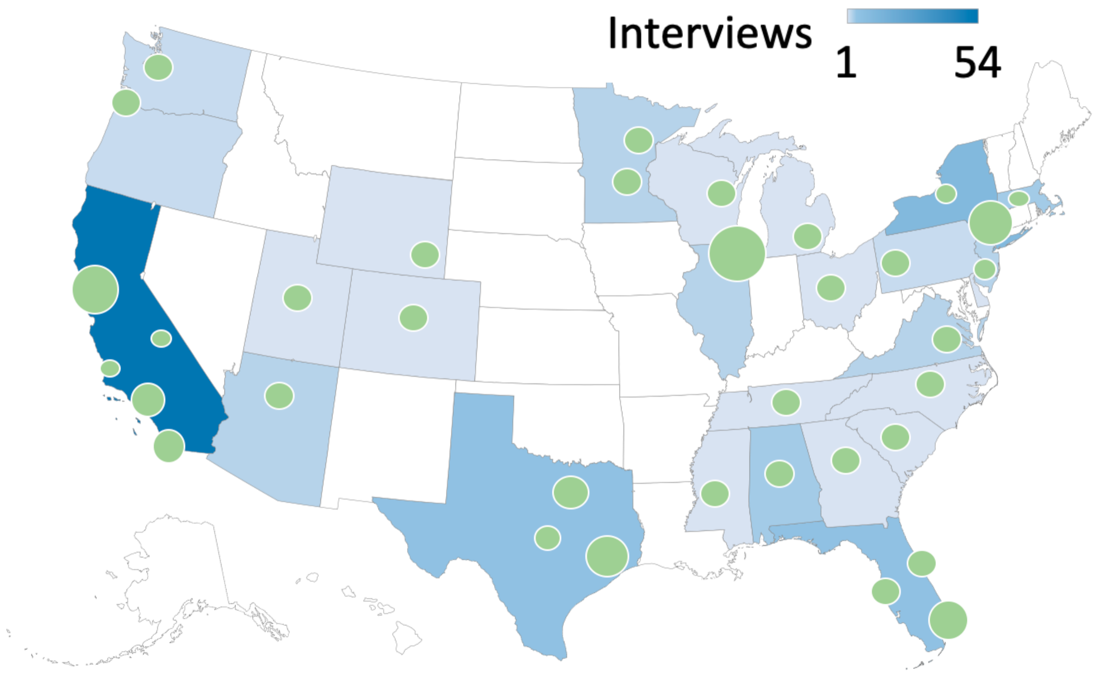

Figure 6.1: A map of our physical therapy interviewees during the NSF I-Corps Program.

The COVID-19 Global Pandemic has caused an unprecedented need for the advancement of telehealth technologies to provide physical rehabilitation care [154]. While number of telehealth sessions skyrocketed due to the constraints of the pandemic, physical therapists were challenged with the loss of hands-on-patient evaluation methods [155, 156]. Moving forward. we can learn from the shortcomings of current telehealth technologies during the pandemic to design better tools and platforms for therapists and patients. Telehealth for physical rehabilitation has many promising affordances as it provides a more encompassing model of care by increasing accessibility and number of patient visits through remote interaction [154]. Yet, for physical therapy to be effectively implemented in telehealth during and beyond the COVID-19 pandemic, current telehealth platforms must incorporate evidence-based movement metrics in a remote setting [157]. This chapter aims to develop and evaluate the feasibility of a machine learning pipeline using solely the motion tracking data of a mass-produced commercial iVR system to predict a user's joint angles and torques during exercises within virtual environments.

6.2 Methods

The data used to train, validate, and test our model was collected from our previous work entitled "OpenButterfly" which was described in chapter four [3]. OpenButterfly examined the experience of 5 users as they performed shoulder rehabilitation in an iVR exergame over the course of two months, with gameplay shown in Figure 6.2. Our target user group consisted of outpatients recovering from shoulder injuries who failed to continue their at-home exercises and still possessed limited strength and ROM.

Five users participated in two exercise mocap sessions per week within our lab in collaboration with two physical therapists. In total, we collected training data on seven exercises:

- Shoulder Rotation (SR)

Figure 6.2: A participant is shown playing Project Butterfly using the HTC Vive. The silver dots on the player's upper body are the reflective markers of the motion tracking system, and the blue strap on the arm is a wrist weight to help increase strength. The right-hand image is a capture from gameplay. The participant protects the moving butterfly, outlined in green, by placing the blue orb over the butterfly to protect it from the incoming crystals indicated by the yellow arrow.

- Side Arm Raise (SAR)
- Forward Arm Raise (FAR)
- External Rotation (ExR)
- Abducted Rotation (AbR)
- Mixed Press (MxdPr)
- Mixed Circles (MxdCr)

These exercise movements can be seen in Fig. 6.3 where the path of the butterfly is shown and this is what the player attempts to follow. SR, SAR, FAR, ExR, and AbR are all single plane movements that are common rehabilitation exercises while MxdPr and MxdCr are multi-planar movements meant to help the subject actively

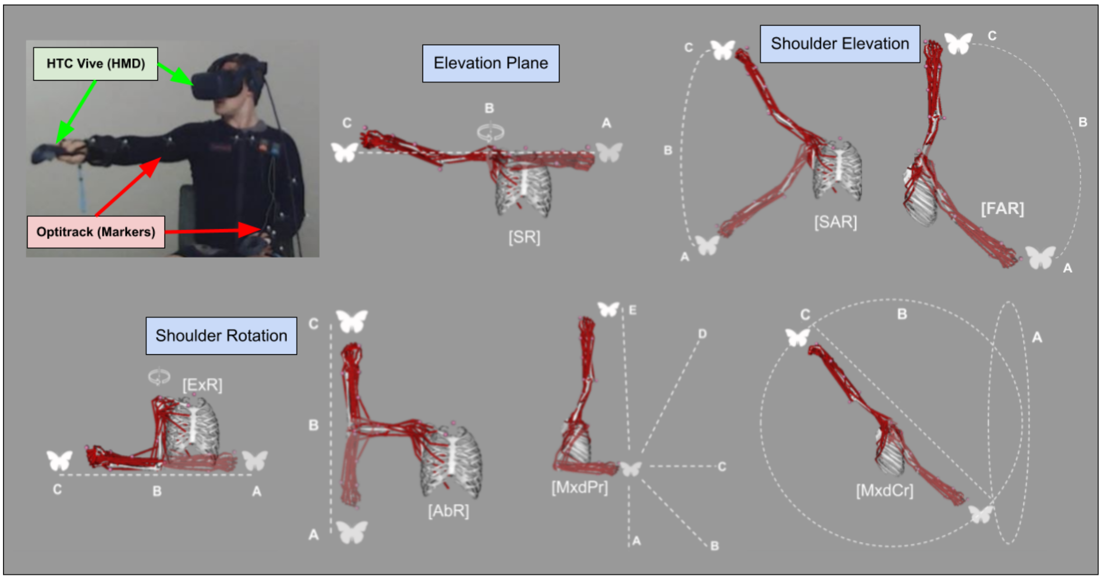

Figure 6.3: Participants played OpenButterfly while seated with ten motion tracking markers placed on bony landmarks as shown in the top left. The game incorporated the seven exercises shown. The dotted line indicates the flight of the butterfly within the game that the users followed with the controller. Letters A-E indicate the direction of the movement. The top row of movements was focused on strength and played with a wrist weight as participants progressed through the protocol. The bottom row of exercises was focused on stretching and was played without weight. The three movements that describe shoulder motion are in the blue text boxes. SR is primarily an Elevation Plane movement, FAR and SAR are primarily Shoulder Elevation movements, and ExR and AbR are primarily Shoulder Rotation movements.

stretch. The first four weeks consisted of games incorporating the movements SR, SAR, and FAR with each exercise performed three times. The following four weeks incorporated four new movements ExR, AbR, MxdPr, MxdCr. During this second phase, all exercises were performed twice during each session. A weighted arm strap was placed on the user's wrist for the exercises SR, SAR, and FAR with increasing weight over the eight week testing period. Our user testing protocol followed this outline for gathering motion capture and iVR tracking data, where sessions lasted a total of 30-45 minutes:

1. Ten reflective markers were placed on bony landmarks of the user.
2. A static T-pose was collected at the beginning of the session for scaling the model.
3. The user was seated, and the headset and controllers were then placed on the user.
4. The user then completed 60 seconds of gameplay followed by 90 seconds of rest. The step was repeated for all exercises for the protocol.

In total, we collected 540 gameplay captures of exercise movement at 60 seconds each.

6.2.1 Motion Capture and Biomechanical Simulation

Optical motion capture systems are considered the gold standard for accuracy and precision [158], yet these types of systems are expensive and often restricted to laboratory environments [159]. With this consideration, we utilized optical motion capture to collect accurate training data from biomechanical simulation (see the top half of Figure 6.4). To collect the training data, we employed eight Optitrack 13W cameras to record ten reflective markers at 120 Hz during gameplay to capture the user's movements [145]. These marker positions are used as input into OpenSim for the Inverse Kinematics Tool, incorporating the upper body model created by Saul et al. [160]. The Inverse Kinematics Tool positions the model to best fit the motion tracking marker data at each time frame. This is done by finding the model pose which minimizes the sum of weighted squared errors of the markers, as shown in Equation 1:

$$\mathrm{SE} = \sum_{i \in m} w_i \left\| x_i^{\mathrm{exp}} - x_i \right\|^2 + \sum_{j \in \mathrm{uc}} w_j (q_j^{\mathrm{exp}} - q_j)^2 \tag{6.1}$$

where

SE is the squared error;

m are the set of markers;

uc are the set of unprescribed coordinates;

x_i^{exp} is the experimental position of marker i;

x_i is the position of the corresponding model marker;

q_j^{exp} is the experimental value for coordinate j;

q_j is the model value for coordinate j;

w_i are the marker weights;

w_j are the coordinate weights.

$q_j = q_j^{\mathrm{exp}}$ for all prescribed coordinates j;

To determine the net forces and torques at each joint, we employ the Inverse Dynamics Tool which uses results from the inverse kinematics and external loads applied to the model. Specifically, OpenButterfly was designed to examine the shoulder joint;

therefore, we focus our model training and prediction on this joint. Below are the classical equations of motion that the Inverse Dynamics Tool uses:

$$M(q)\ddot{q} + C(q, \dot{q}) + G(q) = \tau \tag{6.2}$$

where

$q, \dot{q}, \ddot{q} \in \mathbb{R}^N$ are the vectors of generalized position, velocities, and accelerations, respectively;

$M(q) \in \mathbb{R}^{NxN}$ is the system mass matrix;

$C(q, \dot{q}) \in \mathbb{R}^N$ is the vector of Coriolis and centrifugal forces;

$G(q) \in \mathbb{R}^N$ is the vector of gravitational forces;

$\tau \in \mathbb{R}^N$ is the vector of generalized forces.

The model's motion is defined by the generalized positions, velocities, and accelerations to solve for a vector of generalized forces.

6.2.2 Data Analysis

6.2.2.1 Input Data

During gameplay, the user is seated, and there is minimal movement of the torso or head so the headset moves very little. Additionally, the non-injured arm was not used during gameplay. Therefore, this controller and the headset did not provide valuable input in determining the player's joint mechanics and dynamics of the moving

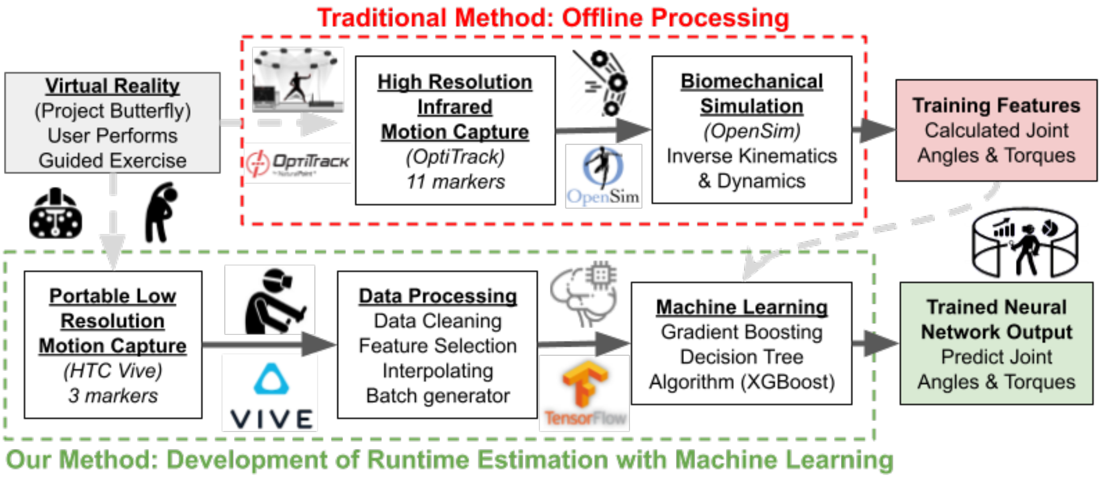

Figure 6.4: Overview of methods to collect data [3], run simulations, and train model. Red pathway shows standard OpenSim method to generate kinematics and dynamics. The green pathway shows our steps to train XGBoost models for predicting the OpenSim results.

arm. Our input features were then the x,y, and z positions along with roll, pitch, and yaw rotation of the moving controller as well as the weight of the arm strap. In total, there were 540 game trials, each recorded for 60 seconds at 120 Hz generating a data set of approximately 3.89 million instances (arm positions). We set aside a set of 54 (10%) randomly selected trials as a test set to test the final models. The remaining 60 second recordings were split into segments of 3 seconds. These shorter segments were used to prevent the model from learning patterns in the movements since some of the movements were repetitive. Each segment was then randomly placed into the training or validation set such that the overall data was split into 80% training, 10% validation, and 10% test.

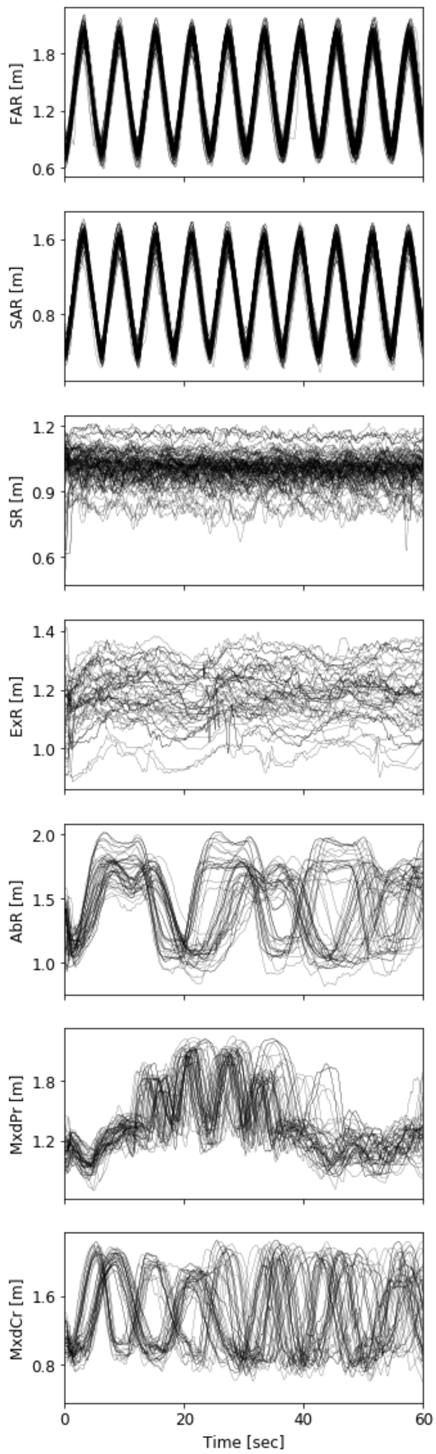

Figure 6.5: Vertical displacement of gameplay controller during each exercise for all users.

6.2.2.2 Machine Learning Model and Prediction

There are many types of machine learning algorithms available that each utilize different types of data and prediction methods. Typically, these algorithms perform regression, clustering, visualization, or classification and can use probabilistic methods, rule-based learners, linear models (e.g. neural networks or support vector machines), decision trees, instance-based learners, or a combination of these [95, 96]. There are pros and cons to each and there is no universal best method for all data sets [97]. Instead, the type of input data needs to be taken into consideration, determine what type of prediction is needed (e.g. binary classification, mulitclass classification, regression, ect.), identify the types of models that are available, and finally consider the pros and cons of those models. Some elements to consider with models are accuracy, interpretability, complexity, scalability, time to train and test, prediction time after training, and generalizability [98, 99, 100, 101, 102].

The input and output data is already known, numeric, and there are multiple input variables making our algorithm selection a supervised multiple regression algorithm. Linear regression and decision trees are commonly used algorithms for these types of tasks. A decision tree is a very simple predictive model that has evolved in the machine learning community through many iterative steps including bagging, random forest, boosting, and gradient boosting [103, 104, 105, 106, 107, 108, 109]. Extreme Gradient Boosting (XGBoost) builds upon all of these methods and has been one of the most widely used machine learning algorithms since being presented at a conference in 2016 out of the University of Washington due to its speed and performance [110]. We opted to use this algorithm because of its ability to accurately train on our type of data as well as its built in regularization methods (LASSO and Ridge) to make sure our models didn't overfit the data.

Six models were trained to produce joint and torque predictions for elevation plane, shoulder elevation, and shoulder rotation as seen in Table 6.1. Shoulder elevation describes rotation about the horizontal axis of the glenohumeral joint, elevation plane describes rotation about the vertical axis of the glenohumeral joint, and shoulder rotation describes rotation about the longitudinal axis of the humerus. The biomechanical simulation data needed to be interpolated to match the collection frequency of the iVR system. The number of estimators was set to 5,000 and the max depth to 10 as values higher than this provided little if any improvement. To prevent overfitting, early stopping rounds were used for each model, so if the model did not improve within five epochs, the training would stop and use the best model. Afterward, the models were used to predict outputs from the unseen test set. The model outputs were then filtered using a 3rd order low-pass Butterworth filter with a cutoff frequency of 3 Hz to remove noise from the signal that is not attributed to the player's movement.

Model Inputs [N=7]	**Model Outputs [N=6]**
Controller Position_x (m)	Elevation Plane Angle (°)
Controller Position_y (m)	Shoulder Elevation Angle (°)
Controller Position_z (m)	Shoulder Rotation Angle (°)
Controller Rotation_x (°)	Elevation Plane Torque (Nm)
Controller Rotation_y (°)	Shoulder Elevation Torque (Nm)
Controller Rotation_z (°)	Shoulder Rotation Torque (Nm)
Arm Strap Weight (kg)	

Table 6.1: Data elements for the machine learning predictive model.

6.2.2.3 Model Evaluation

MAE was used to compare each model's prediction to OpenSim's result within the unseen test set.

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n} \tag{6.3}$$

where

n is number of data points;

y is the prediction of the model;

x is the value obtained from OpenSim.

Motion	OpenSim Angle (°)	OpenSim Torque (Nm)
Elevation Plane	76.9 ±30.3	0.88 ±1.62
Shoulder Elevation	-23.8 ±48.6	9.16 ±2.37
Shoulder Rotation	35.8 ±28.5	101.9 ±246.9

Table 6.2: Mean and standard deviation for OpenSim results from unseen test data set that machine learning models are trying to predict.

Motion	Kinematics MAE (°)	Dynamics MAE (Nm)
Elevation Plane	0.78	0.06
Shoulder Elevation	0.65	0.07
Shoulder Rotation	0.43	2.34

Table 6.3: Mean absolute error between model prediction and OpenSim results for each model's using the unseen test set.

6.3 Results

The motion capture data from OptiTrack was used to generate joint angles and torques in OpenSim. The raw vertical displacement of the controller can be seen for each exercise of all users in Fig. 6.5 and illustrates the different uniformity for each exercise among users. The averages and standard deviations of joint angles and torques of OpenSim can be seen in Table 6.2. Six models were trained to predict joint angles and torques and Fig. 6.6 shows the loss of each of model during training and validation, with early stopping ensuring the models did not over-train. Examples of these results can also be found in Fig. 6.8, Fig. 6.9, and Fig. 6.10.

The MAE comparing the OpenSim results and machine learning models for the unseen test data set is shown in Table 6.3. Based on examining 1000 trails, we found that our trained model can generate predictions in runtime at an average rate of 0.74 ms (+/- 0.36 ms) for a single instance of inputs. An example of two models compared to their corresponding OpenSim outputs can also be seen in Fig. 6.8 for an entire exercise game of 60 seconds. Additional comparisons are illustrated in Fig. 6.9 and Fig. 6.10 for all six models on randomly selected 10-second windows from the unseen test set. Absolute error is also included on the figures to help show the difference between the OpenSim results and model predictions.

6.4 Discussion

This study examined the feasibility and performance of a method for estimating shoulder joint angle and torque from gameplay with an off-the-shelf iVR system. In examining the model performance, the MAE was found to be less than 0.78 degrees for joint angles and less than 2.34 Nm for joint torques indicates that the motion of the iVR system provides enough input for accurate prediction using the XGBoost algorithm. Specifically, the controller's rotation and position, along with the trained arm's wrist weight, are the only metrics needed. This high-accuracy prediction is likely because OpenButterfly was played while seated, so there is minimal torso movement to generate noise.

Subsequently, our results find that iVR systems paired with XGBoost can match or exceed accuracy of the previously mentioned studies in the related works (MAE

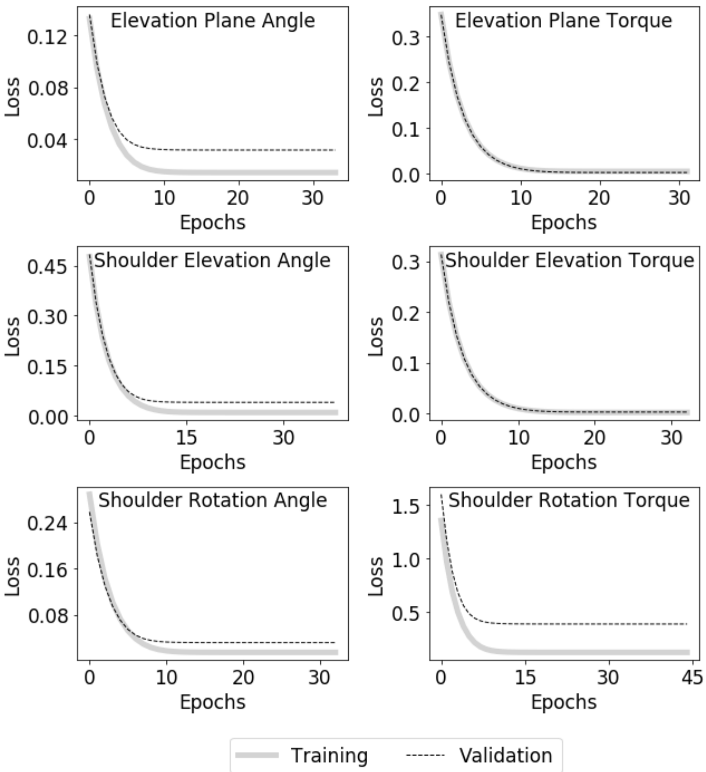

Figure 6.6: Loss function for each model during training to show early stopping preventing over-training.

ranging from 0.8 degrees to 8 degrees for stretch sensor and IMU methods) using an off-the-shelf headset. This is particularly exciting as the widespread adoption of consumer iVR headsets might also be translated for telehealth, potentially utilizing these findings to alleviate the loss of in-person evaluation methods through remote estimation of ROM and joint torques.

Accurate and consistent measurement of ROM is critical to monitoring recovery during physical therapy. Measuring upper limb kinematics is one of the most challenging problems in human motion estimation. The shoulders structure allows for tri-planar movement that cannot be estimated by simple single plane joint models [161, 162, 163]. Our method helps address this complex problem with a low-cost solution that can be used both in the lab and at a patient's home. Unlike prior studies, our approach illustrates that off-the-shelf iVR headsets can be employed for motion analysis in comparison to the complex IMU-based or optical motion capture methods, which require accurate placement on limbs typically dependent on anatomical landmarks [164]. This means that patients can provide more frequent measurements from their homes enabling therapists to have a more detailed remote patient analysis in guiding physical rehabilitation. This technology empowers patients by allowing them to complete at-home guided exercises at a time that works with their schedule over a longer duration and has been shown to aid in recovery over two months [3]. Additionally, our method can provide dynamic measurements as opposed to static ROM measurements so therapists can monitor smoothness of movement quality as well [165]. These measurements can be provided in real-time as the models can generate predictions at a rate of 0.74 ms, potentially enabling synchronous exercise sessions and analysis for physical therapists. Such metrics could be integrated into dashboards for therapist and patient review, as shown in Fig. 6.7, or even used for auto-populating assessment documentation. Real-time metrics can also help with patient safety as the therapist can monitor for incorrect postures, over rotations, and excessive torques to ensure the patient moves their limbs within safe limitations. While our training method uses expensive state-of-the-art motion capture systems and research-grade biomechanical simulation software, none of this is needed on the therapist's or patient's end with our trained model. The model alone can provide these estimations for users performing physical therapy rehabilitation exercises in iVR from games to other virtual experiences.

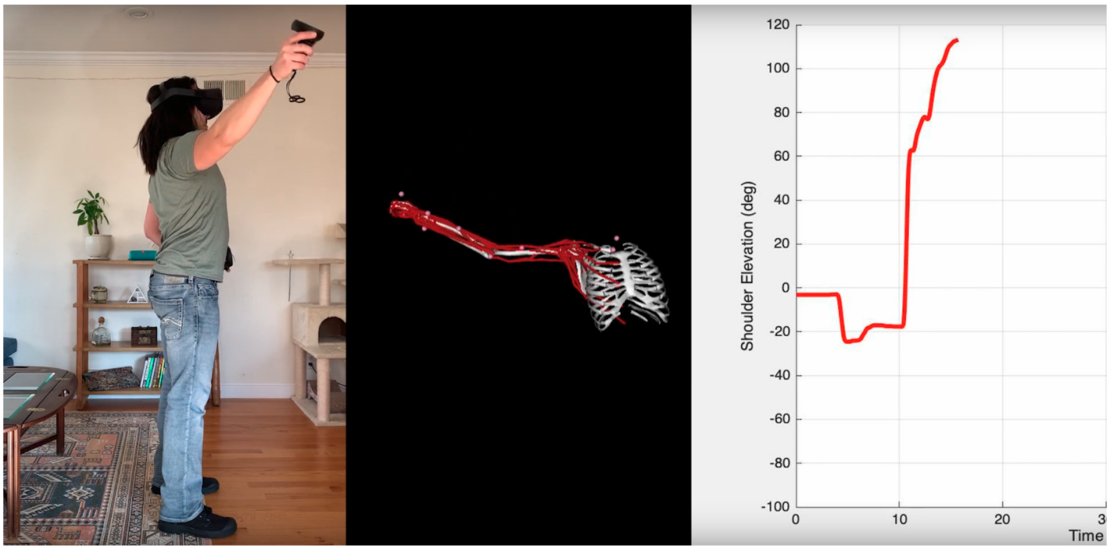

Figure 6.7: We envision as a user plays our games [left image] the avatar can be skeleton showing the users movements [center image] and a dashboard showing the kinematics and dynamics can be running [right image] to show the therapist the relevant metrics needed for remote evaluation.

6.4.1 Limitations of the Study and Future Work

As with any study, there are several limitations that we must consider, many of which could be addressed with future in-person studies. First, the sample of participants to generate training data was small, and each was at similar points in their recovery from shoulder injuries. Future work should have a more diverse user group to train the model to account for the variation of capabilities among users. More users would also allow us to split the data based on the subject so that we can be sure that the algorithm generalizes to unseen users. Second, only seven exercises were examined, six of which were single plane movements. More multi-planar movements should be included in the training data to account for any safe ROM used while playing iVR games. Third,

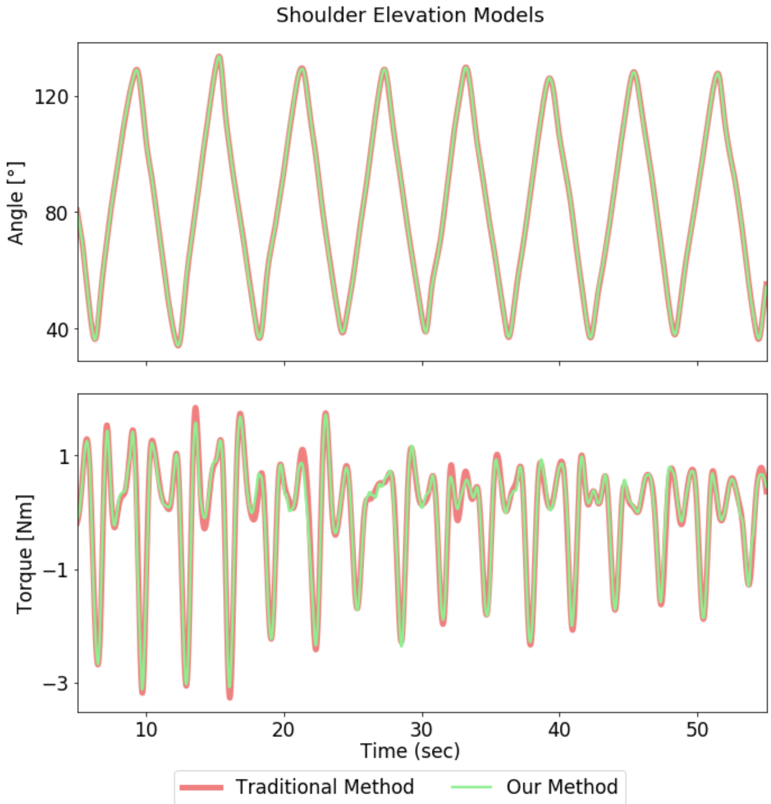

Figure 6.8: An example of OpenSim results and machine learning model predictions for an FAR exercise.

Predictive Mobility Metrics of Seven Exercises through ROM [°]

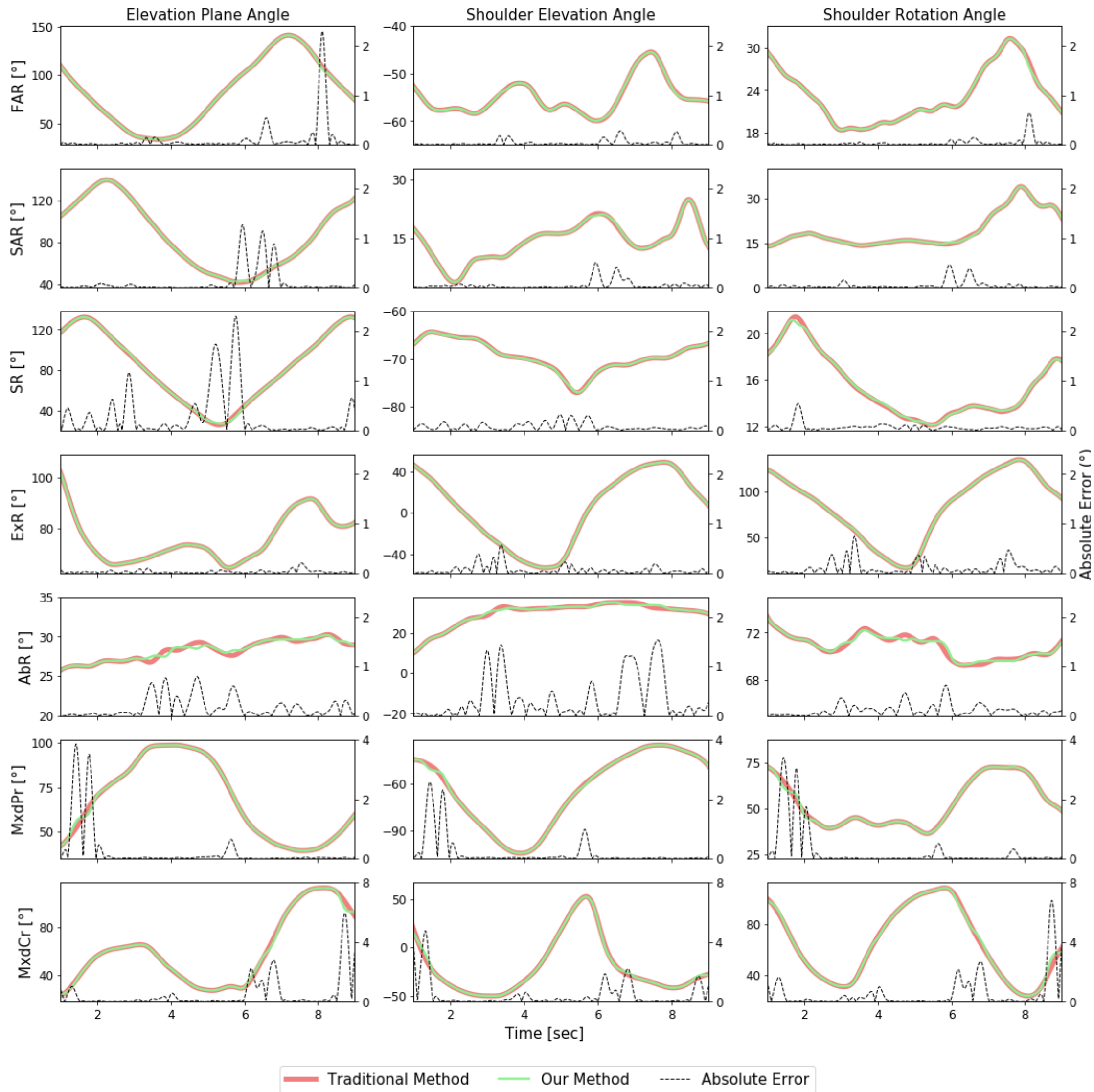

Figure 6.9: Randomly selected segments from the test data set showing the outputs from the traditional method and our method for joint angles for each model with an example for each exercises. Additionally, the absolute error is shown to help see the difference between each method. Exercises are visually demonstrated in Fig. 6.3.

Predictive Strength Metrics of Seven Exercises through Joint Torques [Nm]

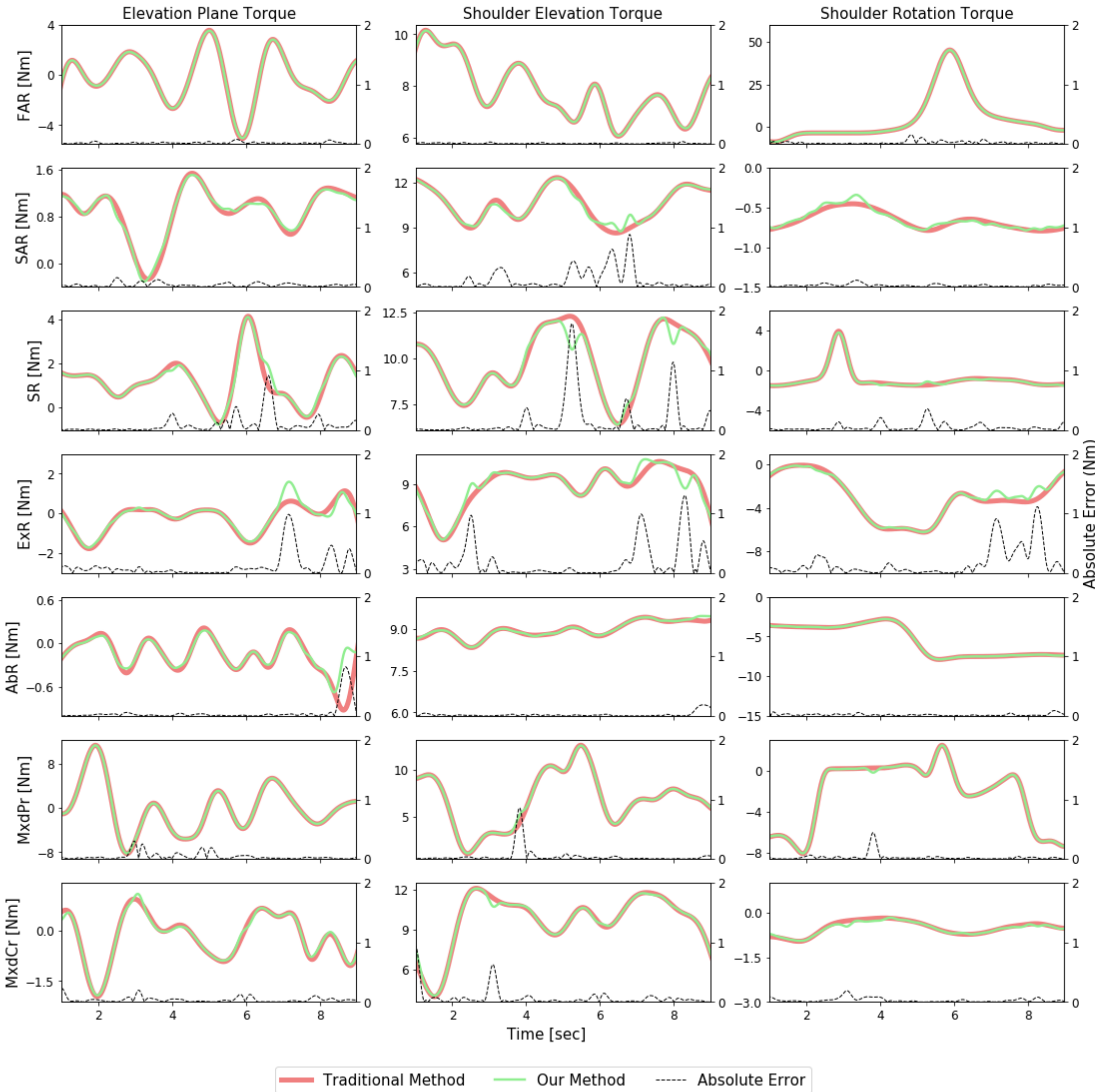

Figure 6.10: Randomly selected segments from the test data set showing the outputs from the traditional method and our method for joint torques for each model with an example for each exercises. Additionally, the absolute error is shown to help see the difference between each method. Exercises are visually demonstrated in Fig. 6.3.

participants played while seated, and the games produced minimal torso movements. Other games that require participants to do movements like stepping, squatting, or bending at the waist should further be examined for validating and extending this model to account for lower-extremity movements. Lastly, only shoulder kinematics and dynamics were examined. Physical therapists of other specializations would benefit with systems that could measure other joints including the elbow, wrist, hip, knee, and ankle. This will likely require input from additional sensor peripherals such as extra controllers placed on the body or computer vision techniques.

Another consideration is to explore differing populations such as those with disabilities (e.g. stroke survivors or cognitive disabilities). Our lab has worked with disability groups within our lab exploring various virtual reality mediums for users with cognitive disabilities, testing soft exo-suits meant for post-stroke rehabilitation, and physical rehabilitation games for users with cognitive disabilities [23, 166, 22, 136]. In the future we will collect motion data of these varying groups to develop more inclusive patient models.

In this work we presented the results using one machine learning algorithm, XGBoost. While it performed well in the future we will do a comparison among other algorithms including Artificial Neural Networks (ANNs), Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM), and Random Forests. In this future work we can compare accuracy of models as well as complexity and training time.

Our lab also aims to make reactive virtual environments by monitoring physiological responses during gameplay using biosensors [144, 3]. Emotion is a crucial component to learning, motivation, interest, and attention during rehabilitation. If we can create a rehabilitation experience that adapts to the user's current emotional state we believe we can improve their experience and outcome. These are future goals we are excited to incorporate into the machine learning model presented in this chapter.

This work suggests that off-the-shelf consumer head-mounted display systems combined with XGBoost can be used to estimate dynamic joint angles and torques in the home setting to help therapists gather relevant metrics throughout the rehabilitation process. These limitations provide a foundation for creating a more generalizable model and future telehealth solutions to empower physical therapists.

6.5 Conclusion

This chapter demonstrated an effective method for estimating shoulder joint angles and torques in real-time during gamified exercises using a head-mounted display iVR system. This method only uses the controllers and headset of intuitive gaming systems, making it ideal for at-home use since a therapist or expert does not need to be physically present. This has the potential to help therapists remotely evaluate a patient and collect metrics that are often difficult to measure with the limited two dimensional videoconferencing. In closing, we can accurately provide evidence based physical rehabilitation metrics through iVR systems paired with predictive models to redefine telehealth. The machine learning related work presented in this chapter has been submitted to a journal and is currently under review with the title "Predictive

Shoulder Kinematics of Rehabilitation Exercises through Immersive Virtual Reality". Another article outlining telehealth and our 130 interviews is under review for another journal and titled "Understanding a Newfound Virtual Reality in Physical Rehabilitation: A Qualitative Study on Therapist Impressions of Telehealth and Technology Needs Amidst the COVID-19 Pandemic."

Chapter 7

Conclusion

7.1 Summary

The work presented in this dissertation explores the use of iVR for the benefit of physical rehabilitation to help determine if current telehealth issues for physical therapists might be solved using 3D HMD rather than 2D media (computers and phones) and if iVR games might help address adherence independent exercise for patients. From our work, we believe the inherent motion capture system of iVR coupled with biomechanical simulations and machine learning will help therapists remotely guide and evaluate patients. This research produced three published journal articles and two more in review, all aimed at understanding and designing virtual reality rehabilitation experiences.

The first step was to understand the differences between HMD and room-scale virtual reality platforms to see if HMD would be a valid option for patients and therapists. We tested an exercise game with users of mixed abilities on each platform and examined their performance, biometric response, and answers to a questionnaire. With this study, we found that a modern HMD such at the HTC Vive is more engaging and produces better physical exercise performance than the more expensive room-scale CAVE platform. This was great to learn as HMD systems are much more affordable and available. With our platform chosen, we began to work on a virtual reality physical therapy game.

Our lab began collaborating with physical therapists to learn what is needed in a rehabilitation game for it to be safe and effective. We limited our study to shoulder injuries so we could rely on the inherent motion capture system (headset and controllers) and not have to incorporate additional sensors that would make it more difficult for therapists and patients to use. Five subjects recovering from shoulder injuries played our game twice a week for eight weeks to see if their strength and range of motion improved. Through this work we determined that our game helped the subjects further recover from their injuries. We also used biomechanical simulations to determine joint kinematics and dynamics to understand the forces on the body and how they change as a player progresses through their recovery.

From this data, we were also able to examine long-term gaming effects. This rehabilitative game was designed to help with long-term adherence since many patients don't continue to perform their at-home exercises, limiting their recovery. If we can create a rehabilitative game that remains engaging for months rather than days, we can increase the rehabilitation timeline and improve recovery. We learned that our game kept users engaged after the initial novelty effect and this is likely attributed to having users progress from simple to more complex movements.

With our initial rehabilitation game, we were awarded a $50,000 NSF grant to interview 130 physical therapists to understand how we might use our technology to help them and their patients. During the COVID-19 pandemic, most physical therapy clinics either closed down due to social distancing or adapted and used telehealth platforms. We learned that another important use of our technology was to adapt our research to help with accurate remote evaluations. Physical therapists had extreme difficulty getting quantifiable metrics like joint angles and forces using 2D videoconferencing style platforms. Now that HMD iVR systems are less than $300, we believed this was a viable option and we had already shown their performance for rehabilitation. These interviews led to the next step of our research.

Therapists need help with remote evaluations to understand a patient's current capabilities and document their progress for safe rehabilitation. Ee used the OptriTrack motion capture system to record the user's movements during gameplay in our previous work. We then determined inverse kinematics and dynamics using our OpenSim model with the motion capture data. These simulations produce the metrics therapists are looking for, but the motion capture system and simulation program are not feasible tools for therapists or patients. Our goal was to produce the simulation results with only the iVR system and a trained machine learning model. To do this, we used the position and rotation of the headset and controllers as input to train a model that would predict the joint angle and torque results of the OpenSim simulation. XGBoost models were trained and tested, yielding results similar to other standard physical therapy tools such as goniometers. With this work, our lab hopes to continue testing these models and collaborating with physical therapists to produce tools that are needed for more equitable care.

7.2 Future Work

The COVID pandemic has been a challenging year for all and limited our ability to collect in-person data. The number of participants for the physical therapy studies were small and all participants were at a similar points in their recovery process. In future studies there should be a larger and more diverse user groups to train the models to help account for variations among users. This group should also include users with mental and physical disabilities to create a more inclusive model to help therapists perform remote evaluations. Another limitation is the number of exercises. The exercises were selected based on feedback about the participants from therapists collaborating on the project to ensure safety. Future work should incorporate more movements to see if models and generalize to the many possible positions of the arm. Our users played seated and produced minimal torso movement. This worked well for shoulder exercises but other exercise games incorporate core movement which should be examined to see if models can include these types of exercises. Only shoulder kinematics and dynamics were examined since this was the type of injury we focused on for our study. It would be beneficial to see if it is possible to create models to predict whole-body movement so that other joints can be examined. Additionally, with the rapid development of many iVR systems there needs to be testing performed with each system to determine its performance and capabilities. These future directions will be explored by the lab once in-person research begins again so that more participants can help train the models.

We believe that this work will empower therapists and patients alike. Therapists will be able to conduct accurate remote evaluations, guide patients, and monitor rehabilitation. Reliable telehealth platforms will remove barriers such as distance, transportation, and time off of work needed for in-person physical therapy sessions that limits many patients. Removing these barriers can help patients participate in more sessions leading to a more complete recovery. For all of this to happen accurate metrics needed by therapists need to be on platforms accessible to therapists and patients. iVR is becoming more accurate, less expensive, and more accessible making it a great platform to develop machine learning models on. This work has shown the potential for upper body metrics and machine learning methods that can be expanded upon for a more inclusive model and eventually a remote evaluation platform for physical therapists and patients.

7.3 Acknowledgment

This material is based upon work supported by the National Science Foundation under Grant No #1521532. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. The authors thank the student participants that helped make this study possible. The authors would like to thank Dana Seth, PT, DPT, for her guidance on developing progression movements for the game to ensure user safety, and Rebecca Friedman, DSC, for her advice and suggestions for therapist needs. The following UC Santa Cruz students are acknowledged for their support: Teresa Faasolo, Evanjelin Mahmoodi, Conrad Esch, and Nico Hawthorne. Thank you to the participants of this study, without their time and effort, none of this would be possible. I was supported by the Cota-Robles Fellowship and Campus Fellow Award during my research and I am extremely grateful for the opportunities they provided.

Bibliography

[1] H. Moore, *MATLAB for Engineers.* Pearson, 2017.

[2] S. L. Delp, F. C. Anderson, A. S. Arnold, P. Loan, A. Habib, C. T. John, E. Guendelman, and D. G. Thelen, "Opensim: open-source software to create and analyze dynamic simulations of movement," *IEEE transactions on biomedical engineering*, vol. 54, no. 11, pp. 1940–1950, 2007.

[3] M. O. Powell, A. Elor, M. Teodorescu, and S. Kurniawan, "Openbutterfly: Multimodal rehabilitation analysis of immersive virtual reality for physical therapy," *American Journal of Sports Science and Medicine*, vol. 8, no. 1, pp. 23–35, 2020.

[4] J. Diemer, G. W. Alpers, H. M. Peperkorn, Y. Shiban, and A. Mühlberger, "The impact of perception and presence on emotional reactions: a review of research in virtual reality," *Frontiers in psychology*, vol. 6, 2015.

[5] D. Gromala, X. Tong, A. Choo, M. Karamnejad, and C. D. Shaw, "The virtual meditative walk: virtual reality therapy for chronic pain management," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems.* ACM, 2015, pp. 521–524.

[6] H. G. Hoffman, G. T. Chambers, W. J. Meyer, L. L. Arceneaux, W. J. Russell, E. J. Seibel, T. L. Richards, S. R. Sharar, and D. R. Patterson, "Virtual reality as an adjunctive non-pharmacologic analgesic for acute burn pain during medical procedures," *Annals of Behavioral Medicine*, vol. 41, no. 2, pp. 183–191, 2011.

[7] B. O. Rothbaum, M. Price, T. Jovanovic, S. D. Norrholm, M. Gerardi, B. Dunlop, M. Davis, B. Bradley, E. J. Duncan, A. Rizzo *et al.*, "A randomized, double-blind evaluation of d-cycloserine or alprazolam combined with virtual reality exposure therapy for posttraumatic stress disorder in iraq and afghanistan war veterans," *American Journal of Psychiatry*, vol. 171, no. 6, pp. 640–648, 2014.

[8] N. Morina, H. Ijntema, K. Meyerbröker, and P. M. Emmelkamp, "Can virtual reality exposure therapy gains be generalized to real-life? a meta-analysis of studies applying behavioral assessments," *Behaviour research and therapy*, vol. 74, pp. 18–24, 2015.

[9] M. V. Nararro-Haro, H. G. Hoffman, A. Garcia-Palacios, M. Sampaio, W. Alhalabi, K. Hall, and M. Linehan, "The use of virtual reality to facilitate mindfulness skills training in dialectical behavioral therapy for borderline personality disorder: a case study," *Frontiers in psychology*, vol. 7, 2016.

[10] M. Rus-Calafell, J. Gutiérrez-Maldonado, and J. Ribas-Sabaté, "A virtual reality-integrated program for improving social skills in patients with schizophrenia: a pilot study," *Journal of behavior therapy and experimental psychiatry*, vol. 45, no. 1, pp. 81–89, 2014.

[11] Y. Shiban, I. Schelhorn, P. Pauli, and A. Mühlberger, "Effect of combined multiple contexts and multiple stimuli exposure in spider phobia: a randomized clinical trial in virtual reality," *Behaviour research and therapy*, vol. 71, pp. 45–53, 2015.

[12] H. Grillon, F. Riquier, B. Herbelin, and D. Thalmann, "Virtual reality as a therapeutic tool in the confines of social anxiety disorder treatment," *International journal on disability and human development*, vol. 5, no. 3, pp. 243–250, 2006.

[13] D. Corbetta, F. Imeri, and R. Gatti, "Rehabilitation that incorporates virtual reality is more effective than standard rehabilitation for improving walking speed, balance and mobility after stroke: a systematic review," *Journal of physiotherapy*, vol. 61, no. 3, pp. 117–124, 2015.

[14] K. R. Lohse, C. G. Hilderman, K. L. Cheung, S. Tatla, and H. M. Van der Loos, "Virtual reality therapy for adults post-stroke: a systematic review and meta-analysis exploring virtual environments and commercial games in therapy," *PloS one*, vol. 9, no. 3, p. e93318, 2014.

[15] H. Mousavi Hondori and M. Khademi, "A review on technical and clinical impact of microsoft kinect on physical therapy and rehabilitation," *Journal of Medical Engineering*, vol. 2014, 2014.

[16] M. S. Cameirão, S. Bermúdez, and P. Verschure, "Virtual reality based upper extremity rehabilitation following stroke: a review," *Journal of CyberTherapy & Rehabilitation*, vol. 1, no. 1, pp. 63–74, 2008.

[17] G. Saposnik, M. Levin, S. O. R. C. S. W. Group *et al.*, "Virtual reality in stroke rehabilitation," *Stroke*, vol. 42, no. 5, pp. 1380–1386, 2011.

[18] J. Iruthayarajah, A. McIntyre, A. Cotoi, S. Macaluso, and R. Teasell, "The use of virtual reality for balance among individuals with chronic stroke: a systematic review and meta-analysis," *Topics in stroke rehabilitation*, vol. 24, no. 1, pp. 68–79, 2017.

[19] J. Crosbie, S. Lennon, J. Basford, and S. McDonough, "Virtual reality in stroke rehabilitation: still more virtual than real," *Disability and rehabilitation*, vol. 29, no. 14, pp. 1139–1146, 2007.

[20] A. Elor and J. Ward, "Accessibility needs of extended reality hardware: A mixed academic-industry reflection," *Interactions*, vol. 28, no. 3, pp. 42–46, 2021.

[21] M. Beccue and C. Wheelock, "Research report: Virtual reality for consumer markets," Tractica Research, Tech. Rep., Q4 2016. [Online]. Available: https://www.tractica.com/research/virtual-reality-for-consumer-markets/

[22] A. Elor, S. Kurniawan, and M. Teodorescu, "Towards an immersive virtual reality game for smarter post-stroke rehabilitation," in *2018 IEEE International Conference on Smart Computing (SMARTCOMP)*, IEEE. Taormina: IEEE, Jun. 2018, pp. 219–225. [Online]. Available: https://ieeexplore.ieee.org/document/8421352/

[23] A. Elor, M. Teodorescu, and S. Kurniawan, "Project star catcher: A novel

immersive virtual reality experience for upper limb rehabilitation," *ACM Transactions on Accessible Computing (TACCESS)*, vol. 11, no. 4, p. 20, Nov. 2018. [Online]. Available: https://dl.acm.org/doi/10.1145/3265755

[24] G. Riva, F. Morganti, and M. Villamira, "Immersive virtual telepresence: virtual reality meets ehealth," in *Cybertherapy.* IOS Press, 2004, pp. 255–262.

[25] P. Lindner, A. Miloff, S. Fagernäs, J. Andersen, M. Sigeman, G. Andersson, T. Furmark, and P. Carlbring, "Therapist-led and self-led one-session virtual reality exposure therapy for public speaking anxiety with consumer hardware and software: A randomized controlled trial," *Journal of anxiety disorders*, vol. 61, pp. 45–54, 2019.

[26] A. Won, J. Bailey, J. Bailenson, C. Tataru, I. Yoon, and B. Golianu, "Immersive virtual reality for pediatric pain," *Children*, vol. 4, no. 7, p. 52, 2017.

[27] L. Li, F. Yu, D. Shi, J. Shi, Z. Tian, J. Yang, X. Wang, and Q. Jiang, "Application of virtual reality technology in clinical medicine," *American journal of translational research*, vol. 9, no. 9, p. 3867, 2017.

[28] K. E. Laver, B. Lange, S. George, J. E. Deutsch, G. Saposnik, and M. Crotty, "Virtual reality for stroke rehabilitation," *Cochrane database of systematic reviews*, no. 11, 2017.

[29] H. L. Miller and N. L. Bugnariu, "Level of immersion in virtual environments impacts the ability to assess and teach social skills in autism spectrum disorder,"

Cyberpsychology, Behavior, and Social Networking, vol. 19, no. 4, pp. 246–256, 2016.

[30] R. M. Baños, C. Botella, M. Alcañiz, V. Liaño, B. Guerrero, and B. Rey, "Immersion and emotion: their impact on the sense of presence," *CyberPsychology & Behavior*, vol. 7, no. 6, pp. 734–741, 2004.

[31] L. Chittaro, R. Sioni, C. Crescentini, and F. Fabbro, "Mortality salience in virtual reality experiences and its effects on users' attitudes towards risk," *International Journal of Human-Computer Studies*, vol. 101, pp. 10–22, 2017.

[32] A. Elor and A. Song, "isam: Personalizing an artificial intelligence model for emotion with pleasure-arousal-dominance in immersive virtual reality," in *2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)*. IEEE, 2020, pp. 572–576.

[33] A. Elor, A. Song, and S. Kurniawan, "Understanding emotional expression with haptic feedback vest patterns and immersive virtual reality," in *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 2021, pp. 183–188.

[34] R. Soares, E. Siqueira, M. Miura, T. Silva, and C. Castanho, "Biofeedback sensors in game telemetry research," *Simpósio Brasileiro de Jogos e Entretenimento Digital*, pp. 81–89, 2016.

[35] H. D. Critchley, "Electrodermal responses: what happens in the brain," *The Neuroscientist*, vol. 8, no. 2, pp. 132–142, 2002.

[36] W. Boucsein, *Electrodermal activity.* Springer Science & Business Media, 2012.

[37] V. N. Salimpoor, M. Benovoy, G. Longo, J. R. Cooperstock, and R. J. Zatorre, "The rewarding aspects of music listening are related to degree of emotional arousal," *PloS one*, vol. 4, no. 10, p. e7487, 2009.

[38] R. J. Rajae-Joordens, "Measuring experiences in gaming and tv applications," in *Probing experience.* Springer, 2008, pp. 77–90.

[39] M. S. Cameirao, I. B. S. Bermúdez, E. Duarte Oller, and P. F. Verschure, "The rehabilitation gaming system: a review," *Stud Health Technol Inform*, vol. 145, no. 6, 2009.

[40] R. A. Cohen, "Yerkes–dodson law," *Encyclopedia of clinical neuropsychology*, pp. 2737–2738, 2011.

[41] M. Liu, D. Fan, X. Zhang, and X. Gong, "Human emotion recognition based on galvanic skin response signal feature selection and svm," in *2016 International Conference on Smart City and Systems Engineering (ICSCSE).* IEEE, 2016, pp. 157–160.

[42] R. Ramirez and Z. Vamvakousis, "Detecting emotion from eeg signals using the emotive epoc device," in *International Conference on Brain Informatics.* Springer, 2012, pp. 175–184.

[43] A. Al-Nafjan, M. Hosny, Y. Al-Ohali, and A. Al-Wabil, "Review and classification of emotion recognition based on eeg brain-computer interface system research: a systematic review," *Applied Sciences*, vol. 7, no. 12, p. 1239, 2017.

[44] A. Goshvarpour, A. Abbasi, and A. Goshvarpour, "An accurate emotion recognition system using ecg and gsr signals and matching pursuit method," *biomedical journal*, vol. 40, no. 6, pp. 355–368, 2017.

[45] H. Marzbani, H. R. Marateb, and M. Mansourian, "Neurofeedback: a comprehensive review on system design, methodology and clinical applications," *Basic and clinical neuroscience*, vol. 7, no. 2, p. 143, 2016.

[46] J. F. Lubar, "Discourse on the development of eeg diagnostics and biofeedback for attention-deficit/hyperactivity disorders," *Biofeedback and Self-regulation*, vol. 16, no. 3, pp. 201–225, 1991.

[47] J. Lubar, M. Swartwood, J. Swartwood, and D. Timmermann, "Quantitative eeg and auditory event-related potentials in the evaluation of attention-deficit/hyperactivity disorder: Effects of methylphenidate and implications for neurofeedback training," *Journal of Psychoeducational Assessment*, vol. 34, pp. 143–160, 1995.

[48] G. Deuschl, "Recommendations for the practice of clinical neurophysiology," *guidelines of the International Federation of Clinical Neurophysiology*, 1999.

[49] L. F. Haas, "Hans berger (1873–1941), richard caton (1842–1926), and electroen-

cephalography," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 74, no. 1, pp. 9–9, 2003.

[50] R. R. Llinás, "Intrinsic electrical properties of mammalian neurons and cns function: a historical perspective," *Frontiers in cellular neuroscience*, vol. 8, p. 320, 2014.

[51] J. J. Foster, D. W. Sutterer, J. T. Serences, E. K. Vogel, and E. Awh, "Alpha-band oscillations enable spatially and temporally resolved tracking of covert spatial attention," *Psychological science*, vol. 28, no. 7, pp. 929–941, 2017.

[52] M. Rangaswamy, B. Porjesz, D. B. Chorlian, K. Wang, K. A. Jones, L. O. Bauer, J. Rohrbaugh, S. J. O'connor, S. Kuperman, T. Reich *et al.*, "Beta power in the eeg of alcoholics," *Biological psychiatry*, vol. 52, no. 8, pp. 831–842, 2002.

[53] J. Baumeister, T. Barthel, K.-R. Geiss, and M. Weiss, "Influence of phosphatidylserine on cognitive performance and cortical activity after induced stress," *Nutritional neuroscience*, vol. 11, no. 3, pp. 103–110, 2008.

[54] P. M. Walker, *Chambers dictionary of science and technology.* Kingfisher, 1999.

[55] C. Iber and C. Iber, *The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications.* American Academy of Sleep Medicine Westchester, IL, 2007, vol. 1.

[56] J. A. Hobson and E. F. Pace-Schott, "The cognitive neuroscience of sleep: neu-

ronal systems, consciousness and learning," *Nature Reviews Neuroscience*, vol. 3, no. 9, p. 679, 2002.

[57] F. Brigo, "Intermittent rhythmic delta activity patterns," *Epilepsy & Behavior*, vol. 20, no. 2, pp. 254–256, 2011.

[58] J. D. Green and A. A. Arduini, "Hippocampal electrical activity in arousal," *Journal of neurophysiology*, vol. 17, no. 6, pp. 533–557, 1954.

[59] I. Whishaw and C. H. Vanderwolf, "Hippocampal eeg and behavior: change in amplitude and frequency of rsa (theta rhythm) associated with spontaneous and learned movement patterns in rats and cats," *Behavioral biology*, vol. 8, no. 4, pp. 461–484, 1973.

[60] J. O'Keefe and N. Burgess, "Theta activity, virtual navigation and the human hippocampus," *Trends in cognitive sciences*, vol. 3, no. 11, pp. 403–406, 1999.

[61] M. E. Hasselmo and H. Eichenbaum, "Hippocampal mechanisms for the context-dependent retrieval of episodes," *Neural networks*, vol. 18, no. 9, pp. 1172–1190, 2005.

[62] J. R. Hughes, "Gamma, fast, and ultrafast waves of the brain: their relationships with epilepsy and behavior," *Epilepsy & Behavior*, vol. 13, no. 1, pp. 25–31, 2008.

[63] W. Singer and C. M. Gray, "Visual feature integration and the temporal correlation hypothesis," *Annual review of neuroscience*, vol. 18, no. 1, pp. 555–586, 1995.

[64] C. Vanderwolf, "Are neocortical gamma waves related to consciousness?" *Brain research*, vol. 855, no. 2, pp. 217–224, 2000.

[65] E. M. Whitham, T. Lewis, K. J. Pope, S. P. Fitzgibbon, C. R. Clark, S. Loveless, D. DeLosAngeles, A. K. Wallace, M. Broberg, and J. O. Willoughby, "Thinking activates emg in scalp electrical recordings," *Clinical neurophysiology*, vol. 119, no. 5, pp. 1166–1175, 2008.

[66] S. Yuval-Greenberg, O. Tomer, A. S. Keren, I. Nelken, and L. Y. Deouell, "Transient induced gamma-band response in eeg as a manifestation of miniature saccades," *Neuron*, vol. 58, no. 3, pp. 429–441, 2008.

[67] S. O'Nuallain, "Zero power and selflessness: What meditation and conscious perception have in common," *Journal of Cognitive Sciences*, vol. 4, no. 2, pp. 46–64, 2009.

[68] B. of Labor Statistics. (2020) Healthcare physical therapy. [Online]. Available: https://www.bls.gov/ooh/healthcare/physical-therapists.htm

[69] A. S. Tenforde, J. E. Hefner, J. E. Kodish-Wachs, M. A. Iaccarino, and S. Paganoni, "Telehealth in physical medicine and rehabilitation: a narrative review," *PM&R*, vol. 9, no. 5, pp. S51–S58, 2017.

[70] S. L. Grona, B. Bath, A. Busch, T. Rotter, C. Trask, and E. Harrison, "Use of videoconferencing for physical therapy in people with musculoskeletal conditions:

a systematic review," *Journal of telemedicine and telecare*, vol. 24, no. 5, pp. 341–355, 2018.

[71] M. A. Cottrell and T. G. Russell, "Telehealth for musculoskeletal physiotherapy," *Musculoskeletal Science and Practice*, vol. 48, p. 102193, 2020.

[72] S. Mani, S. Sharma, B. Omar, A. Paungmali, and L. Joseph, "Validity and reliability of internet-based physiotherapy assessment for musculoskeletal disorders: a systematic review," *Journal of telemedicine and telecare*, vol. 23, no. 3, pp. 379–391, 2017.

[73] A. Rajagopal, C. L. Dembia, M. S. DeMers, D. D. Delp, J. L. Hicks, and S. L. Delp, "Full-body musculoskeletal model for muscle-driven simulation of human gait," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 10, pp. 2068–2079, 2016.

[74] E. M. Arnold, S. R. Hamner, A. Seth, M. Millard, and S. L. Delp, "How muscle fiber lengths and velocities affect muscle force generation as humans walk and run at different speeds," *Journal of Experimental Biology*, vol. 216, no. 11, pp. 2150–2160, 2013.

[75] T. K. Uchida, J. L. Hicks, C. L. Dembia, and S. L. Delp, "Stretching your energetic budget: how tendon compliance affects the metabolic cost of running," *PloS one*, vol. 11, no. 3, p. e0150378, 2016.

[76] H. Baskar and S. M. R. Nadaradjane, "Minimization of metabolic cost of muscles

based on human exoskeleton modeling: a simulation," *Int. J. Biomed. Eng. Sci*, vol. 3, no. 4, p. 9, 2016.

[77] B. J. de Kruif, E. Schmidhauser, K. S. Stadler, and L. W. O'Sullivan, "Simulation architecture for modelling interaction between user and elbow-articulated exoskeleton," *Journal of Bionic Engineering*, vol. 14, no. 4, pp. 706–715, 2017.

[78] G. S. Virk, U. Haider, I. Nyoman, N. Masud, I. Mamaev, P. Hopfgarten, and B. Hein, "Design of exo-legs exoskeletons," in *ASSISTIVE ROBOTICS: Proceedings of the 18th International Conference on CLAWAR 2015.* World Scientific, 2016, pp. 59–66.

[79] D. S. Pina, A. A. Fernandes, R. N. Jorge, and J. Gabriel, "Designing the mechanical frame of an active exoskeleton for gait assistance," *Advances in Mechanical Engineering*, vol. 10, no. 2, p. 1687814017743664, 2018.

[80] J. A. Dienes, X. Hu, K. D. Janson, C. Slater, E. A. Dooley, G. J. Christ, and S. D. Russell, "Analysis and modeling of rat gait biomechanical deficits in response to volumetric muscle loss injury," *Frontiers in Bioengineering and Biotechnology*, vol. 7, p. 146, 2019.

[81] M. Marieswaran, A. Sikidar, A. Goel, D. Joshi, and D. Kalyanasundaram, "An extended opensim knee model for analysis of strains of connective tissues," *Biomedical engineering online*, vol. 17, no. 1, p. 42, 2018.

[82] Z. F. Lerner, B. C. Gadomski, A. K. Ipson, K. K. Haussler, C. M. Puttlitz, and

R. C. Browning, “Modulating tibiofemoral contact force in the sheep hind limb via treadmill walking: Predictions from an opensim musculoskeletal model,” *Journal of Orthopaedic Research*, vol. 33, no. 8, pp. 1128–1133, 2015.

[83] J. W. Rankin, J. Rubenson, and J. R. Hutchinson, “Inferring muscle functional roles of the ostrich pelvic limb during walking and running using computer optimization,” *Journal of the Royal Society Interface*, vol. 13, no. 118, p. 20160035, 2016.

[84] J. M. Burnfield, K. R. Josephson, C. M. Powers, and L. Z. Rubenstein, “The influence of lower extremity joint torque on gait characteristics in elderly men,” *Archives of physical medicine and rehabilitation*, vol. 81, no. 9, pp. 1153–1157, 2000.

[85] L. Ballaz, M. Raison, C. Detrembleur, G. Gaudet, and M. Lemay, “Joint torque variability and repeatability during cyclic flexion-extension of the elbow,” *BMC sports science, medicine and rehabilitation*, vol. 8, no. 1, p. 8, 2016.

[86] A. K. Gillawat and H. J. Nagarsheth, “Human upper limb joint torque minimization using genetic algorithm,” in *Recent Advances in Mechanical Engineering.* Springer, 2020, pp. 57–70.

[87] K. Kiguchi and Y. Hayashi, “An emg-based control for an upper-limb power-assist exoskeleton robot,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 42, no. 4, pp. 1064–1071, 2012.

[88] D. H. Perrin, R. J. Robertson, and R. L. Ray, "Bilateral isokinetic peak torque, torque acceleration energy, power, and work relationships in athletes and nonathletes," *Journal of Orthopaedic & Sports Physical Therapy*, vol. 9, no. 5, pp. 184–189, 1987.

[89] J. Hamill and K. M. Knutzen, *Biomechanical basis of human movement.* Lippincott Williams & Wilkins, 2006.

[90] M. T. Farrell and H. Herr, "Angular momentum primitives for human turning: Control implications for biped robots," in *Humanoids 2008-8th IEEE-RAS International Conference on Humanoid Robots.* IEEE, 2008, pp. 163–167.

[91] S. M. Bruijn, P. Meyns, I. Jonkers, D. Kaat, and J. Duysens, "Control of angular momentum during walking in children with cerebral palsy," *Research in developmental disabilities*, vol. 32, no. 6, pp. 2860–2866, 2011.

[92] C. Nott, R. R. Neptune, and S. Kautz, "Relationships between frontal-plane angular momentum and clinical balance measures during post-stroke hemiparetic walking," *Gait & posture*, vol. 39, no. 1, pp. 129–134, 2014.

[93] R. R. Neptune and C. P. McGowan, "Muscle contributions to whole-body sagittal plane angular momentum during walking," *Journal of biomechanics*, vol. 44, no. 1, pp. 6–12, 2011.

[94] I. El Naqa and M. J. Murphy, "What is machine learning?" in *machine learning in radiation oncology.* Springer, 2015, pp. 3–11.

[95] A. K. Tanwani, J. Afridi, M. Z. Shafiq, and M. Farooq, "Guidelines to select machine learning scheme for classification of biomedical datasets," in *European Conference on Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics*. Springer, 2009, pp. 128–139.

[96] A. Assareh, M. H. Moradi, and L. G. Volkert, "A hybrid random subspace classifier fusion approach for protein mass spectra classification," in *European Conference on Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics*. Springer, 2008, pp. 1–11.

[97] L. Ward, A. Agrawal, A. Choudhary, and C. Wolverton, "A general-purpose machine learning framework for predicting properties of inorganic materials," *npj Computational Materials*, vol. 2, no. 1, pp. 1–7, 2016.

[98] F. Doshi-Velez and B. Kim, "Towards a rigorous science of interpretable machine learning," *arXiv preprint arXiv:1702.08608*, 2017.

[99] S. García, A. Fernández, J. Luengo, and F. Herrera, "A study of statistical techniques and performance measures for genetics-based machine learning: accuracy and interpretability," *Soft Computing*, vol. 13, no. 10, p. 959, 2009.

[100] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep learning*. MIT press Cambridge, 2016, vol. 1, no. 2.

[101] Y. Yao, Z. Xiao, B. Wang, B. Viswanath, H. Zheng, and B. Y. Zhao, "Com-

plexity vs. performance: empirical analysis of machine learning as a service," in *Proceedings of the 2017 Internet Measurement Conference*, 2017, pp. 384–397.

[102] N. Asadi, J. Lin, and A. P. De Vries, "Runtime optimizations for tree-based machine learning models," *IEEE transactions on Knowledge and Data Engineering*, vol. 26, no. 9, pp. 2281–2292, 2013.

[103] A. Brifcani and A. Issa, "Intrusion detection and attack classifier based on three techniques: a comparative study," *Eng. & Tech. Journal*, vol. 29, no. 2, pp. 368–412, 2011.

[104] J. Liang, Z. Qin, S. Xiao, L. Ou, and X. Lin, "Efficient and secure decision tree classification for cloud-assisted online diagnosis services," *IEEE Transactions on Dependable and Secure Computing*, 2019.

[105] F. J. Yang, "An extended idea about decision trees," in *2019 International Conference on Computational Science and Computational Intelligence (CSCI)*. IEEE, 2019, pp. 349–354.

[106] L. Breiman, "Bagging predictors," *Machine learning*, vol. 24, no. 2, pp. 123–140, 1996.

[107] M. R. Segal, "Machine learning benchmarks and random forest regression," 2004.

[108] R. E. Schapire, "The boosting approach to machine learning: An overview," *Nonlinear estimation and classification*, pp. 149–171, 2003.

[109] A. Natekin and A. Knoll, "Gradient boosting machines, a tutorial," *Frontiers in neurorobotics*, vol. 7, p. 21, 2013.

[110] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785–794.

[111] R. Mitchell and E. Frank, "Accelerating the xgboost algorithm using gpu computing," *PeerJ Computer Science*, vol. 3, p. e127, 2017.

[112] S. Lee, J. K. Kim, X. Zheng, Q. Ho, G. Gibson, and E. P Xing, "On model parallelization and scheduling strategies for distributed machine learning," 2014.

[113] D. K. Srivastava and P. Nair, "Employee attrition analysis using predictive techniques," in *International Conference on Information and Communication Technology for Intelligent Systems.* Springer, 2017, pp. 293–300.

[114] R. Jain and A. Nayyar, "Predicting employee attrition using xgboost machine learning approach," in *2018 International Conference on System Modeling & Advancement in Research Trends (SMART).* IEEE, 2018, pp. 113–120.

[115] M. A. Fauzan and H. Murfi, "The accuracy of xgboost for insurance claim prediction," *Int. J. Adv. Soft Comput. Appl*, vol. 10, no. 2, 2018.

[116] H. Services, "About hope services," H. Services, Ed., 01.24.2020. [Online]. Available: https://www.hopeservices.org/about-hope-services/

[117] Polar, "Polar OH1: optical heart rate sensor," Polar, Ed., 29.06.2019, https://www.polar.com/us-en/products/accessories/oh1-optical-heart-rate-sensor. [Online]. Available: https://www.polar.com/us-en/products/accessories/oh1-optical-heart-rate-sensor

[118] Neulog, "GSR logger sensor NUL-217," Neulog, Ed., 29.06.2019, https://neulog.com/gsr/. [Online]. Available: https://neulog.com/gsr/

[119] InteraXon, "Featured research with muse," M. Research, Ed., 29.06.2019, https://choosemuse.com/muse-research/developer.choosemuse.com/tools/available-data. [Online]. Available: https://choosemuse.com/muse-research/,developer.choosemuse.com/tools/available-data

[120] G. Roelkens, J. V. Campenhout, J. Brouckaert, D. V. Thourhout, R. Baets, P. R. Romeo, P. Regreny, A. Kazmierczak, C. Seassal, X. Letartre, G. Hollinger, J. M. Fedeli, L. D. Cioccio, and C. Lagahe-Blanchard, "III-V/Si photonics by die to wafer bonding," *Materials Today*, vol. 10, no. 7-8, pp. 36–43, 2007.

[121] S. Bhayee, P. Tomaszewski, D. H. Lee, G. Moffat, L. Pino, S. Moreno, and N. A. Farb, "Attentional and affective consequences of technology supported mindfulness training: a randomised, active control, efficacy trial," *BMC psychology*, vol. 4, no. 1, p. 60, 2016.

[122] N. Kovacevic, P. Ritter, W. Tays, S. Moreno, and A. R. McIntosh, "'my virtual

dream' : Collective neurofeedback in an immersive art environment," *PloS one*, vol. 10, no. 7, p. e0130129, 2015.

[123] O. E. Krigolson, C. C. Williams, A. Norton, C. D. Hassall, and F. L. Colino, "Choosing muse: Validation of a low-cost, portable eeg system for erp research," *Frontiers in neuroscience*, vol. 11, p. 109, 2017.

[124] I. NaturalPoint, "NaturalPoint, Inc. DBA OptiTrack Motion Capture System," 29.01.2019. [Online]. Available: https://optitrack.com/

[125] H. Corporation, "HTC Corporation: Vive Pro HMD," 29.01.2019. [Online]. Available: https://www.vive.com/us/product/vive-pro/

[126] C. Jennett, A. L. Cox, P. Cairns, S. Dhoparee, A. Epps, T. Tijs, and A. Walton, "Measuring and defining the experience of immersion in games," *International journal of human-computer studies*, vol. 66, no. 9, pp. 641–661, 2008.

[127] J. D. Gibbons and S. Chakraborti, *Nonparametric statistical inference.* Springer, 2011.

[128] M. Hollander, D. A. Wolfe, and E. Chicken, *Nonparametric statistical methods.* John Wiley & Sons, 2013, vol. 751.

[129] D. J. Higham and N. J. Higham, *MATLAB guide.* Siam, 2016, vol. 150.

[130] B. Rosner, R. J. Glynn, and M.-L. T. Lee, "The wilcoxon signed rank test for paired comparisons of clustered data," *Biometrics*, vol. 62, no. 1, pp. 185–192, 2006.

[131] D. A. Bowman, A. Datey, U. Farooq, Y. Ryu, and O. Vasnaik, "Empirical comparisons of virtual environment displays," Department of Computer Science, Virginia Polytechnic Institute & State . . . , Tech. Rep., 2001.

[132] K. Meyerbröker, N. Morina, G. Kerkhof, P. M. Emmelkamp *et al.*, "Virtual reality exposure treatment of agoraphobia: A comparison of computer automatic virtual environment and head-mounted display." *Annual Review of Cybertherapy and Telemedicine*, vol. 9, no. 1, pp. 41–45, 2011.

[133] K. Kim, M. Z. Rosenthal, D. Zielinski, and R. Brady, "Comparison of desktop, head mounted display, and six wall fully immersive systems using a stressful task," in *2012 IEEE Virtual Reality Workshops (VRW)*. IEEE, 2012, pp. 143–144.

[134] M. Cordeil, T. Dwyer, K. Klein, B. Laha, K. Marriott, and B. H. Thomas, "Immersive collaborative analysis of network connectivity: Cave-style or head-mounted display?" *IEEE transactions on visualization and computer graphics*, vol. 23, no. 1, pp. 441–450, 2017.

[135] A. Philpot, M. Glancy, P. J. Passmore, A. Wood, and B. Fields, "User experience of panoramic video in cave-like and head mounted display viewing conditions," in *Proceedings of the 2017 ACM International Conference on Interactive Experiences for TV and Online Video*, 2017, pp. 65–75.

[136] A. Elor, S. Lessard, M. Teodorescu, and S. Kurniawan, "Project butterfly: Synergizing immersive virtual reality with actuated soft exosuit for upper-

extremity rehabilitation," in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, IEEE. Osaka, Japan: IEEE, Mar. 2019, pp. 1448–1456. [Online]. Available: https://ieeexplore.ieee.org/document/8798014/

[137] F. Jiang, X. Yang, and L. Feng, "Real-time full-body motion reconstruction and recognition for off-the-shelf vr devices," in *Proceedings of the 15th ACM SIGGRAPH Conference on Virtual-Reality Continuum and Its Applications in Industry-Volume 1*. ACM, 2016, pp. 309–318.

[138] A. Inc, "Augmented reality - arkit 3," A. Developer, Ed., 29.06.2019. [Online]. Available: https://developer.apple.com/augmented-reality/arkit/

[139] Orion, "Motion capture, vr, games - project orion," I. Limited, Ed., 29.06.2019. [Online]. Available: https://ikinema.com/orion

[140] J. Earles, R. A. Folen, and L. C. James, "Biofeedback using telemedicine: Clinical applications and case illustrations," *Behavioral Medicine*, vol. 27, no. 2, pp. 77–82, 2001.

[141] I. E. Sutherland, "A head-mounted three dimensional display," in *Proceedings of the December 9-11, 1968, fall joint computer conference, part I*. ACM, 1968, pp. 757–764.

[142] I. Sutherland, "The ultimate display," *Multimedia: From Wagner to virtual reality*, pp. 506–508, 1965.

[143] J. Bown, E. White, and A. Boopalan, "Looking for the ultimate display: A brief

history of virtual reality," in *Boundaries of Self and Reality Online.* Elsevier, 2017, pp. 239–259.

[144] A. Elor, M. Powell, E. Mahmoodi, N. Hawthorne, M. Teodorescu, and S. Kurniawan, "On shooting stars: Comparing cave and hmd immersive virtual reality exergaming for adults with mixed ability," *ACM Transactions on Computing for Healthcare*, vol. 1, no. 4, pp. 1–22, 2020.

[145] I. D. O. NaturalPoint, "Optitrack - industry leading precision motion capture and 3d tracking systems for video game design, animation, virtual reality, robotics, and movement sciences." O. M. C. Systems, Ed., 23.01.2020. [Online]. Available: https://optitrack.com/

[146] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex fourier series," *Mathematics of computation*, vol. 19, no. 90, pp. 297–301, 1965.

[147] Unity Technologies, "Unity real-time development platform — 3d, 2d vr & ar," *Internet: https://unity.com/ [Jun. 06, 2019]*, 2019.

[148] G. Rossum, "Python reference manual," 1995.

[149] W. B. Kibler, J. McMULLEN, and T. Uhl, "Shoulder rehabilitation strategies, guidelines, and practice," *Orthopedic Clinics*, vol. 32, no. 3, pp. 527–538, 2001.

[150] W. B. Kibler, "Shoulder rehabilitation: principles and practice." *Medicine and science in sports and exercise*, vol. 30, no. 4 Suppl, pp. S40–50, 1998.

[151] M. Rosenfield, S. Jahan, K. Nunez, and K. Chan, "Cognitive demand, digital screens and blink rate," *Computers in Human Behavior*, vol. 51, pp. 403–406, 2015.

[152] D.-H. Huang, S.-W. Chou, Y.-L. Chen, and W.-K. Chiou, "Frowning and jaw clenching muscle activity reflects the perception of effort during incremental workload cycling," *Journal of sports science & medicine*, vol. 13, no. 4, p. 921, 2014.

[153] A. Elor, M. O. Powell, E. Mahmoodi, M. Teodorescu, and S. Kurniawan, "Gaming beyond the novelty-effect of immersive virtual reality for physical rehabilitation," *IEEE Transactions on Games*, 2021.

[154] A. C. Lee, "Covid-19 and the advancement of digital physical therapist practice and telehealth," *Physical therapy*, vol. 100, no. 7, pp. 1054–1057, 2020.

[155] J. Wosik, M. Fudim, B. Cameron, Z. F. Gellad, A. Cho, D. Phinney, S. Curtis, M. Roman, E. G. Poon, J. Ferranti *et al.*, "Telehealth transformation: Covid-19 and the rise of virtual care," *Journal of the American Medical Informatics Association*, vol. 27, no. 6, pp. 957–962, 2020.

[156] K. A. Bland, A. Bigaran, K. L. Campbell, M. Trevaskis, and E. M. Zopf, "Exercising in isolation? the role of telehealth in exercise oncology during the covid-19 pandemic and beyond," *Physical Therapy*, vol. 100, no. 10, pp. 1713–1716, 2020.

[157] L. O. Dantas, R. P. G. Barreto, and C. H. J. Ferreira, "Digital physical therapy in the covid-19 pandemic," *Brazilian journal of physical therapy*, 2020.

[158] I. H. Lopez-Nava and A. Muñoz-Meléndez, "Wearable inertial sensors for human motion analysis: A review," *IEEE Sensors Journal*, vol. 16, no. 22, pp. 7821–7834, 2016.

[159] P. Müller, M.-A. Bégin, T. Schauer, and T. Seel, "Alignment-free, self-calibrating elbow angles measurement using inertial sensors," *IEEE journal of biomedical and health informatics*, vol. 21, no. 2, pp. 312–319, 2016.

[160] K. R. Saul, X. Hu, C. M. Goehler, M. E. Vidt, M. Daly, A. Velisar, and W. M. Murray, "Benchmarking of dynamic simulation predictions in two software platforms using an upper limb musculoskeletal model," *Computer methods in biomechanics and biomedical engineering*, vol. 18, no. 13, pp. 1445–1458, 2015.

[161] K. Zhang, P. Werner, M. Sun, F. X. Pi-Sunyer, and C. N. Boozer, "Measurement of human daily physical activity," *Obesity research*, vol. 11, no. 1, pp. 33–40, 2003.

[162] H. Ehsani, M. Poursina, M. Rostami, A. Mousavi, M. Parnianpour, and K. Khalaf, "Efficient embedding of empirically-derived constraints in the ode formulation of multibody systems: Application to the human body musculoskeletal system," *Mechanism and Machine Theory*, vol. 133, pp. 673–690, 2019.

[163] R. Krishnan, N. Björsell, E. M. Gutierrez-Farewik, and C. Smith, "A survey of human shoulder functional kinematic representations," *Medical & biological engineering & computing*, vol. 57, no. 2, pp. 339–367, 2019.

[164] M. Rigoni, S. Gill, S. Babazadeh, O. Elsewaisy, H. Gillies, N. Nguyen, P. N.

Pathirana, and R. Page, “Assessment of shoulder range of motion using a wireless inertial motion capture device—a validation study,” *Sensors*, vol. 19, no. 8, p. 1781, 2019.

[165] R. F. Zulkarnain, G.-Y. Kim, A. Adikrishna, H. P. Hong, Y. J. Kim, and I.-H. Jeon, “Digital data acquisition of shoulder range of motion and arm motion smoothness using kinect v2,” *Journal of Shoulder and Elbow Surgery*, vol. 26, no. 5, pp. 895–901, 2017.

[166] A. Elor and S. Kurniawan, “The Ultimate Display for Physical Rehabilitation: A Bridging Review on Immersive Virtual Reality,” *Frontiers in Virtual Reality*, vol. 1, p. 585993, Nov. 2020. [Online]. Available: https://www.frontiersin.org/articles/10.3389/frvir.2020.585993/full