Title
Development and Evaluation of Intelligent Immersive Virtual Reality Games to Assist Physical Rehabilitation

Permalink
https://escholarship.org/uc/item/50x7h611

Author
Elor, Aviv

Publication Date
2021

UNIVERSITY OF CALIFORNIA
SANTA CRUZ

DEVELOPMENT AND EVALUATION OF INTELLIGENT IMMERSIVE VIRTUAL REALITY GAMES TO ASSIST PHYSICAL REHABILITATION

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

DOCTOR OF PHILOSOPHY

in

COMPUTATIONAL MEDIA

by

Aviv Elor

December 2021

The Dissertation of Aviv Elor
is approved:

Professor Sri Kurniawan, Chair

Professor Steve Whittaker

Professor Nancy Chen

Peter F. Biehl
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

List of Tables

Abstract

Development and Evaluation of Intelligent Immersive Virtual Reality Games to Assist Physical Rehabilitation

by

Aviv Elor

Physical rehabilitation is an intensive process that often holds many challenges for patients such as a lack of engagement, accessibility, and personalization. Immersive media systems enhanced with serious games afford an opportunity to address these challenges. In this dissertation work, we examine a series of case studies on the gamification of various physical therapy methods (Constraint Induced Movement Therapy and Mirror Visual Feedback Therapy) within immersive virtual environments. We also present novel systems towards personalizing immersive virtual reality experiences for greater emotional and physical intelligence. Through this work, we link immersive virtual reality with the concepts of therapy, human behavior theory, and biofeedback, towards the aim of achieving a high-level overview of serious applications with a particular emphasis on physical rehabilitation. We conclude with discussion on a series of large scale studies that uncover therapist needs and long-term usage of immersive virtual reality for physical therapy. The findings of this work contribute to future usages of immersive virtual reality for physical therapy that can assist both patients and therapists in overcoming barriers related to low engagement, lack of time, location inaccessibility, and success measure accuracy.

None of this work would be possible without you. To my family, it is your love and support that drives me forward. To my friends and colleagues, it is your engagement and energy that keeps me running. To my advisors and mentors, it is your encouragement, patience, and guidance that has helped me find my way. I simply couldn't have done it without you. You have forever changed my life, and I will always grateful.

Acknowledgments

Rupturing my tricep ligament in my dominate elbow on the Judo mat is the best thing that ever happened to me – I have never been more appreciative for such an experience. It is this epicenter that has led me to pursue my academic endeavour, but most importantly it brought me closer to all of you. My doctoral advisor, Sri Kurniawan, shared an endless amount of her time in supporting my growth as a researcher, academic, and member of the UC Santa Cruz Community. My undergraduate advisor, Mircea Teodorescu, opened my eyes to the world of academic research and assistive technologies – sparking a passion that eventually led me to pursue this doctoral degree. Steve Whittaker, Leila Takayama, and Nancy Chen shared their wealth of knowledge as members and mentors of my advancement committee, dissertation committee, and senior collaborators throughout my time in the Computational Media department. I also give a huge thanks to my many collaborators who helped spark the joy, passion, and exploration of the work presented in this thesis: Michael Ora Powell, Ash Robbins, Samantha Conde, Asiiah Song, Evanjelin Mahmoodi, Nico Hawthorne, April Sun, Steven Lessard, Tiffany Thang, Ben Hughes, Alison Crosby, Amy Phung, Everardo Gonzalez, Kakani Katija, Steven Haddock, Benjamin Erwin, Eric Martin, Jared Duval, Rutul Thakkar, Delong Du, Kassandra Chin, Sherry Luo, Magy Seif El-Nasr, Michael John, Maxim Kuznetsov, Rohan Jhangiani, Yohei Rosen, Colleen Bosworth, Benedict Paten, David Haussler, Sam Michael, Joel Ward, and the many more who I am actively working with today.

Beyond my academic mentors, my peers have been influential in supporting my doctoral research. I thank Michael Ora Powell for being the best research partner I have ever worked with from the late nights, study deployment, and beyond. My partner

Samantha Conde kept me smiling throughout the most stressful of times, amplified my brainstorming sessions, and was always there whenever I needed a shoulder to lean on. I also thank Conrad Esch, Pavlo Vlastos, Joseph Adamson, Breanna Baltaxe-Admony, Alex Rinaldi, Peter Cottrell, Gordon Keller, Victoria Ly, Erik Jung, Maryam Tebyani, Alex Mazhari, Pattawong Pansodtee, and the many other members of the UCSC community for their engagement, discussion, and energy that helped develop my research.

I also thank the UCSC Computational Media Department for supporting my research and growth as a doctoral student. Thank you to Emily Jackson for always going above and beyond to help advise me on student affairs. I must also acknowledge and thank the many programs that have financially supported my work throughout these past years including the UCSC Eugene Cota-Robles Fellowship, the UC Global and Community Health Wellbeing Fellowship, the ARCS Foundation, CITRIS and the Banatao Institute, and the National Science Foundation. Lastly, I thank my family for always standing by my side and supporting me throughout this doctoral journey: Elana Elor, Gilly Elor, Ronny Elor, Orry Elor, and Amit Elor. Thank you everyone – I simply couldn't have done it without you.

Part I

INTRODUCTION

Chapter 1

Introduction

For the past five years, I have been investigating the efficacy of VR based physical therapy through user-centered design with over four local health organizations across Santa Cruz California, including users with and without neurological, motor, and cognitive disability between numerous virtual, augmented, and mixed reality systems with the the Assistive Sociotechnical Solutions for Individuals with Special needs using Technology (ASSIST) Lab at UC Santa Cruz directed by Professor Sri Kurniawan. Much of the material in this dissertation contains reprints of my disseminated articles which have included collaborations from a variety of UCSC research labs, departments, graduate students, and undergraduate research assistants [12, 13, 14, 15, 2, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28]. My broader contributions of these studies have been following: (1) developing the experimental stimuli through the unity game engine (including 3DUI interaction, game behavior, and runtime data collection), (2) iteratively evaluating user experience testing through a variety of quantitative/qualitative techniques (including user protocol design, experimental evaluation, and qualitative observation/interviews), and (3) leading dissemination of findings (including data analysis on user testing/gameplay and writing research articles as first author). Subsequently, this dissertation presents many of these case studies and concludes by tying these works into the research agenda of an Ultimate Display for Physical Rehabilitation.

1.1 Contribution

I have been very fortunate to be able to conduct research and disseminate the findings of my dissertation work. These works demonstrate novel system contributions from translating therapy theory into iVR with gamification to methodological evaluations of how users interact between varying mediums of immersive media. I've utilized a variety of research skills: user-centered iterative prototyping, logfile system analysis, usability testing, task-based evaluation, focus groups/interview analysis, thematic analysis with video coding, and a variety of statistical methods for determining significance (e.g. Wilcoxon Rank Test, T-Test, ANOVA, etc) and predicting user behavior (e.g. Linear Mixed Models, Proximal Policy Optimization, General Adversarial Imitation Learning, etc). Through my position in the UCSC ASSIST Labs, this dissertation work has been largely interdisciplinary between my collaborators of the Electrical and Computer Engineering Department (DANSER labs on establishing AI methods for biomechanics), Anthropology Department (on ethnographic methods for user-centered design), and Healthcare (Hope Services Santa Cruz, the Cabrillo College Stroke & Disability Learning Center, the Pajaro Valley Community Health Trust, and the Games for Health Journal Early Career Committee).

More specifically, this dissertation provides the following contributions:

- Part I presents a bridging literature review of immersive media centered around rehabilitation healthcare, behavior theory, and biofeedback. We examine seminal works from each of these areas towards immersive virtual reality experience design for physical rehabilitation, while also considering accessibility needs from both a hardware and software perspective. This chapter concludes by presenting a theoretical framework for adopting engagement and immersion theory in designing more intelligent immersive virtual reality physical rehabilitation experiences.

- Part II explores how immersive virtual reality head-mounted display systems can be utilized for serious game design in physical rehabilitation. We examine the iterative design and evaluation of two serious games that translate Constraint Induced Movement Therapy and Mirror Visual Feedback Therapy into exercise game mechanics for stakeholders with and without physical disabilities. Each case study concludes with design considerations for future researchers aiming to adopt immersive media for physical rehabilitation with head-mounted display systems, wearable robotics, and unique stakeholder needs.

- Part III investigates how immersive virtual reality experiences can be personalized for greater emotional intelligence. We demonstrate how affective models (e.g. the Pleasure-Arousal-Dominance model of emotion) can be translated into immersive virtual reality experience design through artificial intelligence and user co-design through a variety of visual and haptic stimuli. These application studies end with design considerations for future researchers interested in utilizing adaptive affect models within their virtual experiences.

- Part IV investigates how immersive virtual reality experiences can be personalized for greater physical intelligence. We demonstrate how deep reinforcement learning can be utilized to create visually assistive agents in exercise games through Proximal Policy Optimization and General Adversarial Imitation Learning. We also examine the design and evaluation of optimized distributed gradient boosted algorithms for predicting key physical rehabilitation success metrics in immersive virtual reality exercise games with off-the-shelf headsets. This chapter concludes with considerations for utilizing immersive virtual reality with machine learning for telehealth and serious game analytics in physical rehabilitation.

- Part V examines immersive virtual reality as exercise games and a tool for physical rehabilitation. Firstly, we explore the efficacy of head-mounted display systems between room scale virtual reality systems for exercise gaming. Secondly, we present a qualitative study highlighting physical rehabilitation clinician impressions and needs for telehealth stemming from 130 interviews at the peak of the COVID-19 pandemic. Third, we investigate the application of immersive virtual reality exercise gaming for shoulder rehabilitation over two months with five participants and follow a variety of therapist recommended success measures from physical, biometric, and game data. This chapter concludes with considerations for future immersive virtual reality exercise games that adopt physical and emotional intelligence design practices in assisting physical rehabilitation.
- Part VI concludes with a key summary on this dissertation work and considerations for future research in immersive virtual reality for physical rehabilitation.

1.2 Intellectual Merit

This dissertation project is based on a series of pilot studies of immersive virtual reality environments for gamified rehabilitation, as discussed in Part II. This research develops towards an iVR experience to provide a remote medium to monitor user pain, discomfort, mobility, and biometrics during a prescribed exercise session. My research has developed the integration of controlled user interaction scenarios and system logfile data that have been utilized towards developing predictive runtime analytics for exercise in immersive virtual reality. In addition, research was performed with local healthcare organizations in Santa Cruz, California, to assess this technology towards providing patient success metrics and exercise interaction through commercial head-mounted display iVR systems. This technology has the potential to positively change a users's physical therapy experience by significantly reducing traditional clinical and insurance costs, enabling remote access for populations of low-socioeconomic backgrounds, alleviating discomfort for patients, and increasing remote recovery insights tenfold. These contributions may provide new paths forward in designing immersive experiences for serious applications in future research.

1.3 Broader Impacts

The broader impacts/potential of this dissertation is the development of intelligent immersive virtual reality environments for remote physical therapy and monitoring. The project's core technology examines biomechanical analysis to facilitate rehabilitation exercise both in the clinic and at the user's home using a head-mounted display and hand trackers. This provides a means for in-patient success metrics and virtual guidance using a remote virtual platform. Such technology may have the ability to enable greater affordability, accessibility, and accuracy of physical therapy for patients and therapists alike. Patient throughput potentially could be doubled through remote visits in virtual environments and automated physical health documentation. In addition, this work focused on assisting marginalized communities in "medical deserts," where patient care is significantly limited by hospital capacity, travel capability, doctors per population, and cost. With remote tools and predictive physical therapy analytics, more individuals can receive access to treatment regardless of socio-economic and demographic background. In the long term, this technology could lower hospital visits, enable therapy clinics to remain open during shelter-in-place periods, decrease cost for patients and clinics alike, and begin detecting exercise needs earlier to manage the pace of recovery by each user. From a toolbox perspective, this work may help inform future computational platforms towards enabling researchers to bolster their own projects by understanding immersive virtual reality user experience for a variety of serious applications in physical rehabilitation.

Chapter 2

The Ultimate Display for Physical Rehabilitation: A Bridging Review on Immersive Virtual Reality

2.1 Summary

Physical rehabilitation is often an intensive process that presents many challenges, including a lack of engagement, accessibility, and personalization. Immersive media systems enhanced with physical and emotional intelligence can address these challenges. This review paper links immersive virtual reality with the concepts of therapy, human behavior, and biofeedback to provide a high-level overview of health applications with a particular emphasis on physical rehabilitation. We examine each of these crucial areas by reviewing some of the most influential published case studies and theories while also considering their limitations. Lastly, we bridge our review by proposing a theoretical framework for future systems that utilizes various synergies between each of these fields.

2.2 Introduction

In 1968, Ivan Sutherland, one of the godfathers of computer graphics, demonstrated the first head-mounted display (HMD) immersive media system to the world: an immersive Virtual Reality (iVR) headset that enabled users to interactively gaze into a three dimensional (3D) virtual environment [29, 30, 31]. Three years before the "Sword of Damocles," Sutherland described his inspiration for the system in what became one of the most influential essays of immersive media: *"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"* [32]. Morbidness aside, this vision of an ultimate display asks if it is possible to create such a computationally adept medium that reality itself could be simulated with physical response. Sutherland's "Sword of Damocles" helped spark a new age of research aimed at answering this question for both academia and industry in the race to build the most immersive displays for interaction within the virtual world [33, 30]. However, this trend was short-lived due to hardware constraints and costs at the time [33].

The past decade has seen explosive growth in this field, with increases in computational power and affordability of digital systems effectively reducing barriers to technological manufacturing, consumer markets, required skills, and organizational needs [34]. In 2019, seven million commercial HMDs were sold and with sales projected to reach 30 million per year by 2023 [35]. This mass consumer adoption has partly been due to a decrease in hardware cost and a corresponding increase in usability. These commercial systems provide a method for conveying 6-DoF information (position and rotation), while also learning from user behavior and movement. From these observations, we argue that the integration of iVR as a medium for guided physical healthcare may offer a cost-effective and more computationally adept option for exercise.

Reflecting back to Sutherland's vision of an ultimate display, we ask: *what would be the ultimate iVR system for physical rehabilitation?* If the sensation of physical reality can be simulated through computation, how might that reality best help the user with exercises and physical rehabilitation? From these questions, we posit that Sutherland's vision of the Ultimate Display requires augmentation to address a key area for healthcare: an intelligent perspective of how to best assist a user. In this paper, we explore these questions by reviewing immersive virtual reality as it intersects the fields of therapy, human behavior, and biofeedback. Immersive media affords a medium for enhancing the therapy and healthcare process. It establishes a mode for understanding human behavior, simulating perception, and providing physical assistance. Biofeedback provides a methodology for evaluating emotional response, a crucial element of mental health that is not often explored in healthcare. Given these key points, the rest of this introduction describes our motivation and goals in undertaking this study.

2.2.1 A need for a more efficacious healthcare medium with physical therapy

Physical inactivity leads to a decline of health, with significant motor degradation, a loss of coordination, movement speed, gait, balance, muscle mass, and cognition [36, 37, 38]. In contrast, the medical benefits of regular physical activity include prevention of motor degradation, stimulation of weight management, and reduction of the risk of heart disease and certain cancers [39]. While traditional rehabilitation has its merits, compliance in performing physical therapy may be limited due to high costs, lack of accessibility, and low education [40, 41, 38, 42, 43]. These exercises also usually lack positive feedback, which is critical in improving compliance with physical therapy protocol [44]. Taking these issues into consideration, some higher-tech initiatives associated with telemedicine, virtual reality, and robotics programs have been found to be more effective in promoting compliance than traditional paper-based and verbal instructions [45, 46, 47]. These higher-tech exercise programs often use sensors to passively monitor a patient's status or to provide feedback so the action can be modified. They may also use actuators to assist the patient in completing the motion [48, 49]. Thus, technology may enable a patient to better follow their physical therapy program, aiding independent recovery and building on the progress made with the therapist. This raises the question of whether virtual environments in the form of iVR might be a suitable technology to address these issues.

Immersive virtual environments and the recent uptake of serious games have immense potential for addressing these issues. The ability to create stimulating programmable immersive environments has been shown to increase therapy compliance, accessibility, and data throughput [50, 42]. Considerable success has been reported in using virtual environments for therapeutic intervention between psychological and physiological research. However, these systems have been mostly constrained due to cost and hardware limitations [33]. For example, early 2000s head-mounted display systems had significant hardware constraints, such as low resolution and low refreshment rates, which led to non-realistic and non-immersive experiences that induced motion sickness [51]. Therefore, at that time, the potential use of immersive displays as a rehabilitation tool was quite limited.

These challenges are no longer as prevalent today: modern iVR systems have advanced technically and can now enhance user immersion through widening the field of view, increasing frame rate, leveraging low-latency motion capture, and providing realistic surround sound. These mediums are becoming ever more mobile and are now a part of the average consumer's entertainment experience [52]. As a result, we argue that now is the right time to consider these display mediums as a possible means of addressing the need for effective, cost-effective healthcare. It may be possible that for iVR to be used as a vehicle to augment healthcare to assist users in recovery by transforming the "fixing people" mentality [53] of traditional rehabilitation into adventures in the virtual world that provide both meaningful enjoyable experiences and restorative exercise.

2.2.2 Review goals

The goal of this paper is to survey the theory, application, and methodology of influential works in the field of immersive media for the purpose of exploring opportunities towards future research, with the ultimate aim of applying these technologies to engage physical rehabilitation. The subsequent sections of this paper provide a discussion of the following topics:

- the current state of academic research in utilizing iVR for physical rehabilitation and health;
- the behavioral theory behind the success of utilizing iVR;
- the applications of biofeedback and incorporating runtime user analysis in virtual environments;

- and bridging potential synergies between each of these areas towards applying them for future research.

This work will provide an overview of iVR for physical rehabilitation and health through an understanding of both past and current academic projects. We aim to provide an informative view on each of our goals as well as offering suggestions for how these concepts may be used to work towards an ultimate display of physical rehabilitation. We believe that this work will be of interest to interdisciplinary researchers at the intersection of immersive media, affective computing, and healthcare intervention.

2.2.3 Scope and limitations

The term VR was coined long before the advent of recent immersive virtual reality (iVR) systems. This has led to differences in how the term “VR” is applied, and these differences can be seen within the existing literature. For the purposes of this review, we define niVR as non-immersive systems that utilize a monitor and allow user interaction through conventional means such as keyboards, mice, or custom controllers [33]. VR systems that provide a head-mounted display (HMD) with a binocular omni-orientation monitor, along with appropriate three-dimensional spatialized sound, are categorized as iVR. Augmented reality (AR) systems employ virtual feedback by allowing the user to see themselves and their surroundings projected virtually onto a screen, usually in a mirror-like fashion [54]. These systems are similar in how they present movement-based tasks with supplementary visual and auditory feedback, but differ in their interaction methods [55].

Our review focuses on iVR systems for physical rehabilitation, health, and games for health. We examine high-impact case studies, meta-reviews, and position papers from academia with an emphasis on research conducted in the past two decades. This paper provides a high-level overview of each of these areas and their implications for healthcare. However, we must acknowledge that immersive media, and many of the other concepts described in this paper, are rapidly changing fields. Many of the academic work and positions discussed in this paper are likely to change in the future as technology advances. With these considerations in mind, this paper provides a snapshot of these research areas from past to present and derives limitations and challenges from such to infer the need for future research in advancing an ultimate display for physical rehabilitation. We start by examining iVR for healthcare and rehabilitation.

2.3 Immersive Virtual Reality and Therapy

In the past two decades, there have been many publications and studies focusing on VR technologies for application in psychotherapy, physiotherapy, and telerehabilitation. Modern iVR technology is commonly known for its impact on enhancing the video gaming paradigm by deepening user involvement and leading to more dedicated interaction [56]. The increased physical demands of these video gaming platforms have garnered interest for their potential in therapy through repetitive and quantifiable learning protocols [57]. Early research suggests that the use of iVR systems is useful for psychological, physical, and telepresence therapy [58, 59].

2.3.1 Psychological therapy applications

Psychological research has seen an increase in the use of iVR due to its ability to simulate realistic and complex situations that are critical to the success of laboratory-based human behavior investigations [60]. Some of these investigations include the successful reduction of pain through the use of stimuli in iVR. This has shown results equivalent to the effects of a powerful analgesic treatment, such as morphine, for burn victim wound treatment [61, 62]. With the immersive capabilities of modern headsets, such as the HTC Vive and Oculus Rift, there has been an increase in studies reporting positive outcomes of iVR exposure therapies for post-traumatic stress disorder [63, 64], borderline personality disorder [65], phobias [66, 67], and schizophrenia [68], as well as many other psychological therapies. This accelerated iVR use in psychological therapy is often attributed to the relationship between increased presence and emotion [64, 69]. Increasing the number of meaningful stimuli that resonate with the users' engagement using iVR is a crucial factor in influencing user behavior and experience [70], and, with the price of computing devices and hardware decreasing, headsets are becoming more popular and immersive in doing so [52, 35]. Thus, immersion through iVR can lead to greater emotional influence on the user and can incite the desired physiological responses by crafting a stimulating and engaging virtual environment [71]. While this work shows great promise, the psychological application of iVR is still largely underdeveloped and lacking in terms of proven beneficial results. Similar results and benefits can also be seen with physical therapy interventions utilizing iVR.

2.3.2 Physiological therapy applications

Traditional forms of physical therapy and rehabilitation are based on therapist observation and judgment; this process can be inaccurate, expensive, and non-timely [42]. Many studies have indicated that iVR can be an effective tool in improving outcomes compared to conventional physical therapy [72]. Environments can be tailored to cue specific movements in real-time through sensory feedback via the vestibular system and mirror imagery to exemplify desired ranges of motion [73]. With the emergence of new immersive multimedia, iVR experiences with sight, sound, and touch can be integrated into rehabilitation. Studies have indicated that iVR intervention is useful in improving a variety of motor impairments, such as hemiparesis caused by Parkinson's disease, multiple sclerosis, cerebral palsy, and stroke [73].

High repetitions of task-oriented exercises are critical for locomotive recovery, and user adherence to therapy protocol is imperative. iVR-based physical rehabilitation can induce adherence to therapy protocol as successfully as (and sometimes better than) human-supervised protocol due to the capabilities of multi-sensory real-time feedback [50]. Games can be used to guide the user in their movements and provide mechanics to reward optimal exercises [50]. Additionally, this multi-sensory, auditory, and visual feedback can further persuade users to exercise harder through increased stimuli. iVR-based physical rehabilitation also allows for increased quantitative feedback for both the user and the therapist. The capacity of modern iVR systems to implement three-dimensional motion tracking serves as an effective way to monitor progress during rehabilitation, allowing healthcare professionals to obtain a more in-depth view of each user's independent recovery [56].

Multiple reviews have been conducted consisting of hundreds of studies through the past decade, and have concluded that niVR is useful for motor rehabilitation [42, 74, 75]. Many of these studies have confirmed that the use of iVR results in significant improvements when compared to traditional forms of therapy [73, 50]. These studies used Kinect, Nintendo Wii, IREX: Immersive Rehabilitation Exercise, Playstation EyeToy, and CAVE, as well as custom-designed systems. For a given treatment time, the majority of these studies suggested that video game-based rehabilitation is more effective than standard rehabilitation [50, 73, 72, 76]. Subsequently, the physical rehabilitation communities have been enthusiastic about the potential to use gaming to motivate post-stroke individuals to perform intensive repetitive task-based therapy. Some games can combine motion capture as a way to track therapy adherence and progress. Despite these promising studies, technology at the time needed to improve in terms motion-tracking accuracy in order to become more effective, reliable, and accessible [77, 42]. The existing research indicates that that more work is needed to continue gaining a deeper understanding of the efficacy of iVR in rehabilitation [74, 78]. These modern iVR headsets open up new opportunities for accessibility and affordability of treatment.

2.3.3 Telerehabilitation applications

Telerehabilitation approaches provide decreased treatment cost, increased access for patients, and more quantifiable data for therapists [79]. There have been various studies confirming the technical feasibility of in-home telerehabilitation, as well as an increase in the efficiency of these services [80]. In these studies, users generally achieve more significant results in rehabilitation due to the increased feedback from the telerehabilitation VR experience [81]. Due to the mobile and computational nature of VR displays, these iVR telerehabilitation studies suggests that the usability and motivation of the rehabilitation treatment for the user can be sustained while reducing work for therapists and costs for patients [82].

2.3.4 Limitations of current studies for iVR rehabilitation

While iVR has shown great promise from these studies, we must establish whether these HMDs and immersive displays are a truly beneficial medium. The cost of HMDs is reducing and commercial adoption is prevalent [52]. However, research into the effectiveness of iVR as a medium for rehabilitation is still inconsistent and is not often verified for reproducibility. An unfortunate commonality between these studies lies in a lack of reporting methodology, small or non-generalizable user sample sizes, not accounting for the novelty effect, and making blunt comparisons in terms of the effectiveness and usability of such systems. For example, in a review by Parsons et al., hundreds of studies addressing virtual reality exposure therapy for phobia and anxiety were reviewed in terms of affective functioning and behavior change. The biggest issue with Parsons's comparative review was a small sample size and a failure to account for the variety of factors that play into VR. The authors argue that Virtual Reality Exposure Therapy (VRET) is a powerful tool for reducing negative symptoms of anxiety, but could not directly calculate demographics, anxiety levels, phobia levels, presence, and immersion between these studies. While curating this review did provide an active snapshot into VRET usage in academia, it is arguable that the data from these studies may have been weak or biased due to the low sample sizes demonstrating positive results and the missing factors of usability for use beyond a single academic study [83]. A study by Jeffrey et al. examined twenty children who received distraction from IV treatment with two controls; iVR HMDs with a racing game as a distraction and a distraction-free treatment case. The results indicated that pain reduction was significant, with a four-fold decrease in facial pain scale responses in cases where iVR was used [84]. This work positively supports the use of iVR HMDs as a medium for pain reduction, but also lacks a large sample size and provides a somewhat biased comparison of iVR. Is it not to be expected that any distraction of pediatric IV placement would reduce pain? Is iVR versus no distraction a fair comparison to the general protocol for pediatric IV placement? What about the usage of a TV, or even an audiobook, against the iVR case? In another review by Rizzo et al., VRET was studied using an immersive display that showed veterans 14 different scenarios involving combat-related PTSD stimuli. In one trial, 45% of users were found to no longer test positive for PTSD after seven sessions of exposure. In another trial, more than 75% no longer tested positive for PTSD after ten sessions. Most users reported liking the VR solution more than traditional exposure therapy [85]. Again, this use of iVR for therapy focused on a small sample size and specific screening techniques, which must be taken into consideration when reviewing the results. Testing for PTSD change in this context only provides a snapshot of VR's effectiveness. Furthermore, the novelty effect (in the sense that the users are not acclimated to the system) may have a significant influence on the result. Given these points, what would happen when users have fully acclimated to this system – is the promise of VRET therapy demonstrated by Parsons, Jeffery, and Rizzo et al. truly generalizable? Ultimately, the answer may lie in the direct need for more iVR rehabilitation studies to evaluate and transparently disseminate results between the iVR and niVR comparative norms. An ultimate display for physical rehabilitation with the ability to simulate almost any reality in instigating therapeutic goals may have much potential, but we must understand the behavioral theory behind iVR as a vehicle for healthcare.

2.4 Immersive Virtual Reality, Behavior, and Perception

As discussed in the previous section, immersive media systems hold vast potential for synergizing the healthcare process. Rehabilitation research, including physical and cognitive work incorporating iVR-based interventions, has been on the rise in recent years. There is now the ability to create programmable immersive experiences that can directly influence human behavior. Conducting conventional therapy in a iVR environment can enable high-fidelity motion capture, telepresence capabilities, and accessible experiences [72, 13]. Through gamification, immersive environments with commercial iVR HMDs, such as the HTC Vive, can be programmed to increase therapy compliance, accessibility, and data throughput by crafting therapeutic goals as game mechanics [13]. However, what drives the success of iVR healthcare intervention? What aspects of behavioral theory can inform an optimal virtual environment that will assist users during their healthcare experiences? This section aims to explore and understand the theory behind the success of using iVR in healthcare.

2.4.1 The benefits of immersion

iVR provides a means of flexible stimuli through immersion for understanding human behavior in controlled environments. Immersion in a virtual environment can be characterized by the sensorimotor contingencies, or the physical interaction capability of a system [86]. It attributes to how well the system may connect a user in iVR through heightened perception and ability to take action, also known as perceptual immersion [87]. This is dependent on the number of motor channels and the range of inputs provided by the system in order to achieve a high fidelity of sensory stimulation [88]. Subsequently, perceptual immersion also opens an opportunity for psychological immersion [87], enabling users to perceive themselves to be enveloped by and a part of the environment [89].

The success of iVR therapeutic intervention is often attributed to the influence of immersion in terms the ability to enhance the relationship between presence and emotion in an engaging experience, and the influence of this on overcoming adversity in task-based objectives [64]. Immersion can be continuously enhanced through improving graphics, multi-modality, and interaction [86]. Strong immersive stimuli through a iVR system, and the ability to provide a feeling of presence and emotion engagement in a virtual world, are key to influencing user behavior [64, 70, 71]. Because of this, iVR can play an essential role in augmenting the physical therapy process through the benefits of immersion as it corresponds to a greater spatial and peripheral awareness [90].

Higher-immersion virtual environments were found to be overwhelmingly positive in treatment response [91]. The detachment from reality that is induced by immersion in a virtual world can reduce discomfort for a user, even as far as minimizing pain when compared to clinical analgesic treatments [61, 62]. For example, one study found that an iVR world of playful snowmen and snowballs may reduce pain as effectively as morphine during burn victims' wound treatment [92]. Increasing the number of stimuli using iVR is a crucial factor in influencing user experience [70]. With iVR systems becoming ever more affordable and accessible, these immersive environments are becoming available to the average consumer [52].

2.4.2 Presence in the virtual environment

Given the benefits of immersion, from task-based guidance in spatial awareness to enabling psychological engagement, it is critical to quantify the effects of presence through immersion. Diemer et al. have suggested that presence is derived from the technological capabilities of the iVR system and is strengthened by the sense of the immersion of a virtual environment [69]. “Presence” can be defined as the state of existing, occurring, and being present in the virtual environment, and it has been extensively modeled and quantified through past research. Schubert et al. have argued that presence has three dimensions: spatial presence, involvement, and realness [93]. These dimensions are often quantified through a preliminary survey and cognitive scenario evaluation. Witmer et al. have argued that presence is cognitive and is manipulated through directing attention and creating a mental representation of an Immersive Virtual Environment (iVE) [94]. Furthermore, Seth et al. 2012 have argued for the introspective predictive coding model of presence, which posits that presence is not limited to iVR but is “a basic property of normal conscious experience” [95]. This argument rests on a continuous prediction of emotional and introspective states, where the perceiver’s reaction to the stimulus is used to identify success. For example, a fear stimulus as can be utilized during the prediction of emotional states, where the user compares the actual introspective state (fear and its systems) with the predicted emotional state (fear). A higher presence indicates successful suppression of the mismatch between the predicted emotional state versus actual emotional state [69]. Thus, if the prediction of the fear stimuli is victorious over the mismatch of the user’s actual reaction, this may indicate that they were happy, rather than in a state of fear (as was predicted). The idea that suppression of information in a VR experience is vital for presence and the inducing emotion is not new and was previously proposed by Schuemie et al. [96]. Seth et al. have emphasized that the prediction of emotional states from stimuli plays a crucial role in enabling an emotional experience [95]. Parsons’ research supports this claim;

presence is regarded as a necessary mediator to allow “real emotions” to be activated by a virtual environment [83]. However, Diemer et al. has cautioned that research has not yet clarified the relationship between presence and emotional experiences in iVR.

Moreover, quantifying presence is still primarily conceptualized through task-based methods (such as subjective ratings, questionnaires, or interviews), all of which are largely qualitative in nature. A debate between many of these presence theories is whether or not emotion is central to modeling presence. For example, Schubert et al.'s "spatial presence" or Slater's "place illusion" do not require emotion as a prerequisite for presence, which is unlike Diemer's hypothesis of emotion connecting presence and immersion. Given that physical health and recovery has been heavily linked to emotional states [97, 98], we consider Diemer at al.'s model of presence. Therefore, to effect presence in a virtual environment, there is a need of quantifying emotion. How does one model emotion in this regard, or even quantify it?

2.4.3 Emotion and virtual environments

Quantifying the human emotional response to media has been the topic of much debate in academia. Paul Ekman, a pioneer of emotion theory, argued that there are six basic emotions: anger, fear, sadness, enjoyment, disgust, and surprise [99]. He argued that there are nine characteristics of emotions: they have universal signals, they are found between animals, they affect the physiological system (such as the nervous system), there are universal events which invoke emotion, there is coherence in emotional response, they have rapid onset, they have a brief duration, they are appraised automatically (subconsciously), and their occurrence is involuntary [99]. Ekman's theory does not dismiss any affective phenomena, but instead organizes them to highlight the distinction, based on previous research (in the fields of evolution, physiology, and psychology) between the field and his previous work. His theory also provides a means of quantifying emotions using these principals; it offers a theoretical framework for constructing empirical studies to understand affective states as well as basic emotions [99]. Ekman's basic emotions were found and identifiable in media such as music [100] and photos [101].

Since the early 2000s, researchers have examined how technology can extend, emulate, and understand human emotion. Rosalind Picard, the pioneer of affective computing, has expanded upon theories such as Ekman's to build systems that understand emotion and can communicate with humans emotionally [102]. This had lead numerous findings and demonstrations of systems that demonstrate discrete models (including Ekman's basic emotion model, appraisals models, dimensional models, circuit models, and component models) for quantifying emotional response [102, 103]. Moreover, numerous machine learning methods have been demonstrated as emotion inference algorithms, such as classification, artificial neural networks, support vector machines, k-nearest neighbor, decision trees, random forests, naive Bayes, deep learning, and various clustering algorithms [103].

With Diemer et al.'s model, emotional engagement may enhance presence to assist the user in an iVE task. Thus it is useful to quantify a user's emotional response in an iVE. Many studies have examined sense signals and classified patterns as an emotional response from the Autonomic Nervous System. In relation to the basic emotions, Collet et al. have observed patterns in skin conductance, potential, resistance, blood flow, temperature, and instantaneous respiratory frequency through the use of six emotion-inducing slides presented to 30 users in random order [101]. Through the use of questionnaires, Meuleman et al. found that appraisal theory induced the highest emotional response with the HTC Vive iVR System [104]. Liu et al. have utilized real-time EEG-based emotion recognition by applying an arousal-valence emotion model with fractal dimension analysis [105] with 95% accuracy along with the National Institute of Mental Health's (NIMH) Center for Study of Emotion and Attention (CSEA) International Affective Picture System (IAPs) [8]. One of the most widely used metrics for emotion evaluation is the NIMH CSEA Self-Assessment Manikin (SAM) [7]. Waltemate used SAM to evaluate emotion concerning the sense of presence and immersion in embedded user avatars with 3D scans through an iVR social experience [106]. SAM enables the evaluation of dimensional emotion (through quantifying valence, arousal, and dominance) by using a picture-matching survey to evaluate varying stimuli. It has been validated for pictures, audio, words, event-related potentials, functional magnetic resonance imaging, pupil dilation, and more [7, 8, 107, 108].

In addressing emotional experiences that influence presence, or a user's sense of "being in" an iVE, we must consider what influences these experiences. Broadly, the majority of research turns to human perception to answer this question. Previous psychological research on threat perception, fear, and exposure therapy implies a relationship between perception and emotion. Perception influences emotion and presence in an iVE, which enables a controlled environment for identifying the most relevant aspects of each user's emotional experience [70]. The association between perception and conceptual information in iVR must also be considered, as this can play a crucial role in eliciting emotional reactions. For behavior research focusing on areas such as fear, anxiety, and exposure effects, it is vital that iVR is able to induce emotional reactions leading presence and immersion [69]. This can be achieved by adjusting perceptual feedback of a user's actions through visual cues, sounds, touch, and smell to trigger an emotional reaction. This goes two ways, in the sense that iVR allows the consideration of how perception can be influenced by iVR itself while also enabling emotional engagement. Therefore, researchers can dissociate perceptual and informational processes as controlled conditions to manipulate their studies in unique ways using iVR [70]. Given that researchers have found ways to model and influence perception for presence and emotion, what has been done in iVR?

2.4.4 Human perception and multi-sensory displays

Human perception appears to be the ultimate driver of user behavior. Yee et al.'s Proteus effect has demonstrated how both self-representation and context in a virtual environment can be successfully influenced via iVR HMDs [109]. The way we perceive the world around us—through our expectations, self-representation, and situational context—may influence how we act and how we approach behavioral tasks. Human perception is reliant on multimedia sensing, such as processing sight, sound, feel, smell, and taste [110]. This is problematic because the majority of published research on iVR does not account for this; many studies focus on a singular modality such as a sight or sound, and only occasionally connect sight, sound, and feel. However, with modern advances in commercially available hardware, all senses except for taste have the potential to be controlled in a virtual environment.

2.4.4.1 Stimuli and perception

Exploring new input modalities for iVR in physical rehabilitation may help discover new and effective approaches for treatment experience. For example, there have been many studies that have examined how haptic feedback can communicate, help recognize, and inform pattern design for emotions. Bailenson et al. examined how interpersonal touch may reflect emotional expression and recognition through a hand-based force-feedback haptic joystick [111]. They found that users were able to both recognize and communicate emotions beyond chance through the haptic joystick. In a study by Mazzoni et al., the design and evaluation of a haptic glove for mapping emotions evoked by music were found to reliably convey pleasure and arousal [112]. Bonnet et al. found that facial expression emotion recognition was improved when utilizing a "visio-haptic" platform for virtual avatars and a haptic arm joystick [113]. Salminen et al. examined the patterns of a friction-based horizontally rotating fingertip stimulator for pleasure, arousal, approachability, and dominance for hundreds of different stimuli pairs [114]. Fingertip actuation indicated that a change in the direction and frequency of the haptic stimulation led to significantly different emotional information. Obrist et al. demonstrated that patterns in an array of mid-air haptic hand stimulators map onto emotions through varying spatial, directional, and haptic parameters [115]. Miri et al. examined the design and evaluation of vibrotactile actuation patterns for breath pacing to reduce user anxiety [116]. The authors found that frequency, position, and personalization are critical aspects of haptic interventions for social-emotional applications.

Many prior studies have also found that olfactory echoing principle of universal emotions. Fox has examined the human sense of smell and its relationship to taste, human variation, children, emotion, mood, perception, attraction, technology, and related research [117]. Sense of smell is often dependent on age (younger people outperform older people), culture (western cultures differ from eastern cultures), and sex (women outperform men). However, other studies suggest that sense of smell mainly depends on a person's state of mental and physical health, regardless of other factors. Some 80-year-olds have the same olfactory prowess as 20-year-olds, and a study from the University of Pennsylvania showed that people who are blind do not necessarily have a keener sense of smell than sighted people [117]. It appears to be possible to "train" one's sense of smell to be more sensitive. This poses a problem for researchers, as some subjects in repetitive experiments become skilled in this (i.e., the weight of scent differ for people depending on their sensitivity). Subsequently, Fox has argued that "the perception of smell consists not only of the sensation of the odors themselves but of the experiences and emotions associated with sensations" [117]. These smells can evoke strong emotional reactions based on likes and dislikes determined by the emotional association. This occurs because the olfactory system is directly connected with an ancient and primitive part of the brain called the limbic system where only cognitive recognition occurs. Thus, a scent may be associated with the triggering of deeper emotional responses. Similar to the Proteus effect [109], our expectations of an odor influence our perception and mood when encountering the stimulus [117].

In terms of perception, positive emotions are indicated with pleasant fragrances and can affect the perception of other people (such as attractiveness of perfume and photographs). Unpleasant smells tend to lead to more negative emotions and task-based ratings (such as when viewing a picture or a completing survey of pleasant or unpleasant odors). General preferences for smells exist (i.e., that the smell of flowers is pleasant and that the smell of gasoline or body odor is unpleasant). Some fragrances, such as vanilla, are universally perceived as pleasant (which is why most perfumes use vanilla). Perfume makers have also shown that appropriate use of color can better identify our liking of fragrance [117]. This is supported by the work of Hirsch et al., who explored how olfactory aromas can be quantified to demonstrate arousal [118]. They explored 30 different scents via wearable odor masks with 31 male volunteers. By measuring penile blood flow, the authors found that every smell produced an increase of penile blood flow when compared to no odor, and that pumpkin pie and lavender (which, according to Fox, is considered a universally pleasant scent) produced the most blood flow, with a 40% increase [118]. There appear to be universal smells that are coherent across different demographics, similar to Ekman's argument for universal emotions shared by different races, animals, and sexes [117, 99]. An ultimate display that could utilize these smells and adapt to each user's individual preferences by understanding their presence and emotion could be useful in both eliciting an engaging medium of therapy and discovering new universal stimuli.

2.4.4.2 On multi-modal immersive virtual reality environments

Many researchers have started to recognize and explore the potential of multi-modality iVR interfaces. In an exploratory study by Biocca, Kim, and Choi, the authors concluded that presence may derive from multi-modal integration, such as haptic displays, to improve user experiences [119]. Bernard et al. showcased an Arduino-driven haptic suit for astronauts to increase embodied situation awareness, but no evaluation was reported [120]. Goedschalk et al. examined the potential of the commercially available KorFX vest to augment aggressive avatars, but found an insignificant difference between the haptic and non-haptic conditions [121]. And, Krogmeier et al. demonstrated how a bHaptics Tactisuit vest can influence greater arousal, presence, and embodiment in iVR through a virtual avatar "bump" [122]. The authors found significantly greater embodiment and arousal with full vest actuation compared to no actuation. However, this study only examined a singular pattern and one set stimuli.

Numerous examples can also be seen with thermal actuation, haptic retargeting, and olfactory input. For example, Wolf et al. and Peiris et al. explored thermal actuation embedded in iVR HMD facial masks and tangibles which increased enjoyment, presence, and immersion [123, 124]. Doukakis et al. evaluated a modern system for audio-visual-olfactory resource allocation with tri-modal virtual environments which suggested that visual stimuli is the most preferred for low resource scenarios and aural/olfacotry stimuli preference increases significantly when budgeting is available [125]. Warnock et al. have found that multi-modality notifications through visual, auditory, tactile, and olfactory interfaces were significant in personalizing the needs and preferences of home-care tasks for older adults with and without disability [126, 127]. Azmandian et al. used haptic virtual objects to "hack" real-world presence by shifting the coordinates of the virtual world, leading users to believe that three tangible cubes lay on a table when in reality there was only one cube [128]. Olfactory inputs have been found to be incredibly powerful in increasing immersion and emotional response, such as in Ischner et al.'s Brain and Behavioral Laboratory Immersive Oder System [129], Aiken et al. 's review of olfaction for PTSD treatment [130], and Schweizer's application of iVR and olfactory input for training emergency response [131]. Dihn et al. demonstrated that multi-sensory stimuli for an iVR virtual office space can increase both presence and spatial memory from a between-subjects factorial user study that varied level of visual, olfactory, auditory, and tactile information [132]. These systems have shown great promise in personalizing systems with the capability to rapidly adapt to smells in an iVR environment. Beyond these theories and proposed systems, there are many limitations and challenges to keep in mind when translating these theories into applied environments.

2.4.5 Limitations of current studies for iVR behavior and perception

Immersion, presence, and emotion are critical in influencing an engaging, motivating, and beneficial iVR therapy. However, these themes are not analyzed in iVR therapy studies as standard. This may be primarily due to a lack of uniform quantification of these areas. However, there are many surveys and sensing techniques used to quantify biofeedback, such as the NIMH CSEA SAM and valence-arousal models. Even when studies incorporate such considerations, sample sizes are usually small and methodology is not always transparent. A gold standard can be seen with the NIMH CSEA Self-Assessment Manikin [107, 108], for which affect is validated using a stimuli database that has been pre-validated by hundreds of participants. There may be a clear benefit in releasing the iVR stimuli evaluated through the ultimate display to create an international affective database for cross-modal virtual reality stimuli.

The user's understanding of how to perform therapy exercises, as well as their commitment to performing them for the duration of the therapy, is critical to ensure effectiveness of rehabilitation. The emotional response generated by an immersive experience influences user engagement and may motivate patients to continue with the objectives of the virtual experience [71]. Therefore, we ask: *how might we quantify the success of iVR stimuli towards affecting a users emotional engagement?* This leads us to the next section, in which we discuss how understanding the increasing availability of biometric sensors and biofeedback devices for public use may help us find answers to these questions [133].

2.5 Immersive Virtual Reality and Biofeedback

This section aims to identify the theory and usage of biofeedback through a variety of sensory modalities for immersive media and behavioral theory. Biofeedback devices have gained increasing popularity, as they use sensors to gather useful, quantifiable information about user response. For example, the impedance of the sweat glands, or galvanic skin response (GSR), has been correlated to physiological arousal [134, 135]. This activity can be measured through readily available commercial GSR sensors, and has been explored by researchers to measure the arousal created by media such as television, music, and gaming [136, 137]. Different types of iVR media may affect biofeedback performance. Cameiro et al. analyzed niVR-based physical therapy that uses biofeedback to adapt to stroke patients based on the Yerkes-Dodson law [138] or the optimal relationship between task-based performance and arousal [139]. By combining heart rate (HR) with GSR, game events and difficulty were quantitatively measured for each user to evaluate optimal performance. Another example can be seen in the work of Liu et al., in which GSR alone achieved a 66% average emotion classification accuracy for users watching movies [140]. Combined with GSR, HR can indicate the intensity of physical activity that has occurred. There is definite potential in evaluating the GSR and HR of each user to determine the intensity of the stimuli using different systems of iVR. However, GSR and HR are not the only biometric inputs that could be potentially leveraged when understanding an immersive experience.

In another biofeedback modality, commercially available electroencephalography (EEG) sensors have shown great promise in capturing brain activity and even in inferring emotional states [141]. Brain-computer Interfaces (BCI) incorporating EEG devices have become ever more affordable and user-friendly, with computational techniques for understanding user engagement and intent in medical, entertainment, education, gaming, and more [142]. Based on a review of over 280 BCI-related articles, Al-Nafjan et al. have argued that EEG-based emotion detection is experiencing booming growth due to advances in wireless EEG devices and computational data analysis techniques such as machine learning [142]. Accessible and low-cost BCIs are becoming more widely available and accurate in the context of both medical and non-medical applications. They can be used for emotion and intent recognition in entertainment, education, and gaming [142]. When compared with 12 other biofeedback experiments, studies that used EEG alone were able to reach 80% max recognition [143]. Arguably, the most considerable challenges of BCI are costs, the impedance of sensors, data transfer errors or inconsistency, and ease of use [142, 143].

Even with these challenges, EEG has been successfully used to as a treatment tool for understanding conditions like attention deficit/hyperactivity disorder (ADHD), anxiety disorders, epilepsy, and autism [144]. Brain signals that are characteristic of these conditions can be analyzed with EEG biofeedback to serve as a helpful diagnostic and training tool. Sensing apparatus can be coupled with interactive computer programs or wearables to monitor and provide feedback in many situations. By monitoring levels of alertness in terms of average spectral power, EEG can aid in diagnosing syndromes and conditions like ADHD, anxiety, and stroke [145]. Lubar et al. used the brainwave frequency power of game events to extract information about reactions to a repeated auditory stimulus, and have demonstrated significant differences between ADHD and non-ADHD groups [146]. Through exploring different placements and brainwave frequencies of EEG sensors across a user's scalp, different wavebands can be used to infer the emotional state and effect of audio-visual stimuli [147]. For example, Ramirez et al.

used the alpha and beta bands to infer arousal and valence, respectively, which are then mapped to a two-dimensional emotion estimation model [141]. With these examples in mind, how does one quantify brainwaves for emotional inference?

2.5.1 Brainwaves as a means of studying emotional intelligence

Hans Berger, a founding father of EEG, was one of the first to analyze these frequency bands of brain activity and correlate them to human function [148, 149]. The analysis of different brainwave frequencies has been correlated to different psychological functions, such as the 8-13 Hz Alpha band relating to stress [150], the 13-32 Hz Beta band relating to focus [151, 152], the 0.5-4 Hz Delta band relating to awareness [153, 154, 155, 156], the 4-8 Hz Theta band relating to sensorimotor processing [157, 158, 159, 160], and the Gamma band of 32-100 Hz related to cognition [161, 162, 163]. These different frequencies may prove fruitful in quantifying the effects of virtual stimuli during iVR based physical therapy, taking into account the fact that signals may be noisy due to other biological artifacts and must be handled carefully [164, 165, 166]. For example, alpha activity is reduced with open eyes, drowsiness, and sleep [150]; increases in beta waves have been suggesting for active, busy, or anxious thinking and concentration [152]; delta activity spikes with memory foundation [155] such as flashbacks and dreaming [156]; theta activity increases when planning motor behavior [**?**] path spatialization [159] memory, and learning [160]; and gamma shows patterns related to deep thought, consciousness, and meditation [161].

Additionally, there are many methods for evaluating and classifying emotions with brainwaves. Eimer et al. used high-resolution EEG sensing to analyze the processing of Ekman's six basic emotions via facial expression during P300 event-related potential analysis (ERP) [167]. Emotional faces had significantly different reaction times from neutral faces (supporting the rapid onset of emotion Ekman's principle). The authors concluded that ERP facial expression effects gated by spatial attention appear inconsistent, however, ERP effects are directly due to Amygdala activation, they also conclude that ERP results demonstrate facial attention is strongly dependent on facial expression, and that the six basic facial expressions with emotions were strikingly similar [167]. ERPs are an effective way to quantify EEG brainwave readings for emotional analysis, but they are not always reliable. However, they can accurately gauge from an arousal response by looking at a P300 window of revealing stimuli. These techniques open opportunities for estimating emotion through multiple biofeedback modalities.

Researchers have combined these EEG interfaces with other forms of multimodal biometric data collection such as GSR and HR to increase the inference of affective response. By combining GSR with HR and EEG, researchers have been able to increase the accuracy of emotion recognition [143, 140]. Other niVR based games have successfully incorporated the use of these biofeedback markers to determine physiological response [138, 133]. However, there is a lack of studies exploring these biometrics with iVR and physical therapy, such as the one described in this paper. This is particularly true in the case of examining long-term use beyond the novelty period and allowing for user acclimatization to the experimental environment. With such limitations in mind, it is possible that these effects and psychological responses could be quantitatively measured through combining active EEG sensing with the flexible stimuli of iVR gameplay. In the light of this, what has been done to bridge biofeedback to iVR?

2.5.2 Biofeedback systems utilized with virtual reality

The closest experience (albeit not immersive) to the proposed ultimate display augmentation for rehabilitation discussed in this paper can be seen in i Badia et al.'s work on a procedural biofeedback-driven nonlinear 3D-generated maze that utilized the NIMH CSEA International Affective Picture System. VR mental health treatment has seen extensive exploration and promising results over the past two decades. However, most of the experiences are not personalized for treatment, and more personalized treatment is likely to lead to more successful rehabilitation. i Badia has argued for the use of biofeedback strategies to infer the internal state of the patient state [168]. Users navigated a maze where the visuals and music were adapted according to emotional state [168]. The framework incorporated the Unity3D game engine in a procedural content generation through three modules of real-time affective state estimation, event trigger computation, and virtual procedural scenarios. These were connected in a closed-loop during runtime through biofeedback, emotion game events, and sensing trigger events. The software architecture uses any iVR medium and runs the Unity application with a separate process for data acquisition via UDP protocol, which was published and shared as a Unity plugin [168]. Overall results indicated significance for anger, fear, sadness, and neutral (in Friedman analysis), and a Self-Assessment Manikin Indicated significant feelings of pleasantness associated with the experience. However, the game was not explored using an immersive medium (instead, a Samsung TV was used), varying intensity was not explored, and control factors were random to each user, which may have influenced results [168].

Immersive experiences exploring low-cost commercial biofeedback devices have been also been presented, although methodology has not been fully disseminated. Redd et al. found that cancer patients during Magnetic Resonance Imaging responded with a 63% decrease in anxiety with heliotropin (a vanillalike scent) with humidified air when compared to a odorless humidified air alone [169].) Expanding upon this work, Amores et al. utilized a low-cost commercial EEG device, a brain-sensing headband named Muse 2 [170], with an olfactory necklace and immersive virtual reality for promoting relaxation [171]. By programming odor to react to alpha and theta EEG activity within iVR, users demonstrated increases of 25% physiological response and reported relaxation when compared to no stimulis. This may validate the effectiveness of combining iVR with olfactory input, as well as the ability to quantify mental state through physiological changes through low-cost, low-resolution commercial EEG.

In another example, Abdessalem et al. compared mental activity of EEG recordings to the International Affective Picture System for a serious game named "AmbuRun." Users entered an iVR game in which they had to carry a patient in an ambulance to the hospital and drive it through traffic. They evaluated the game with 20 participants, and the difficulty adapted to each user so that higher frustration led to more traffic [172]. The authors identified significant results; 70% of players reported that the game was harder when they were frustrated, while only 15% said they did not notice any change in difficulty. However, this study does not share baseline EEG activity results, nor does it explain the adaptive difficulty algorithms that were used [172]. Other examples realting biofeedback and iVR can be found in the work of Marin et al., who examined EEG and heart rate variability with portable iVR HMDs to elicit emotions by exploring 3D architectural worlds [173]. Kronert et al. developed a custom headband that recorded BVP, PPG, and GSR while adults completed various games in learning environments [174]. Van Rooij developed a game that displayed diaphragmatic breathing patterns in children with the aim of reducing in-game anxiety, and was able to get users to reverse panic attacks [175]. Again, while all these results were highly promising in incorporating biofeedback techniques to augment iVR user experiences, they were also lacking in many areas.

2.5.3 Limitations of current studies for iVR biofeedback

A large amount of work has been done independently in the biofeedback field in terms of methods of sensing mental activity, and there is now a plethora of sensing methods. Some games have been created incorporating biofeedback with promising results. However, these studies are often vague and do not publish stimuli or demos beyond what is written in the paper. In this literature review, we have found that most of these biofeedback games are not multi-modal sensing and thus do not account for any low-resolution sensing or movement artifacts from gameplay through sensor fusion (i.e., HR and GSR could be used with in-game behavior to cross-validate physiological signal change during therapy with EEG sensing). Additionally, the majority of these studies do not incorporate runtime feedback from the user themselves (beyond pre- or post-test surveys). Quantifying emotion is usually done either solely through biofeedback and emotion estimation, or post-test surveys, but never both during runtime. It is possible that biofeedback emotional estimation combined with embedded gameplay surveys may be a way to better objectively measure presence, as long as immersion is not broken when queried for survey response.

Additionally, these studies are often not conducted with multi-modal stimuli. Human perception is inherently multi-modal, and perhaps emotional response may become more accurate when utilizing multiple human senses beyond audio and visual stimuli. What happens when we factor in smell and touch while collecting biofeedback measures within iVR? As with the other limitations discussed in the previous two sections, much of this work is not disseminated beyond the papers themselves (with the exception of i Badia et al.'s published biofeedback plugin [168]. Future researchers can address these limitations by fully disseminating their methodology and algorithms in their work, and such aspects should be transparent towards the design and evaluation of immersive media with biofeedback.

2.6 An Ultimate Display for Physical Rehabilitation

We dedicate this section to expand upon the current literature review and bridge the discussions in the previous sections on immersive virtual reality, rehabilitation, behavioral theory, and biofeedback. In the previous sections, we discussed how the newfound commercial adoption of iVR devices and the affordability of biofeedback devices may lead to new opportunities for adaptive experiences in healthcare that are feasible for the average consumer. iVR-based therapy from psychological, physiological, and telepresence applications have shown great promise and great potential. The theory and success behind iVR as a medium for healthcare intervention is driven by immersion and its relationship with presence and emotion. Because presence and emotion tend to be subjective, quantification of their measures is not always reproducible. However, many quantification methods exist, ranging from a sensing algorithmic approach to a variety of validated surveys. The current literature review has found that more work must be done to provide clear guidelines, universal iVR stimuli to evaluate affect, and an environment that factors multi-modal sensing and stimulation for presence and emotion. These items may address a need for a controllable multi-modal immersive display that can factor in physical and emotional intelligence through both qualitative and quantitative biofeedback.

2.6.1 Augmenting the Ultimate Display

To bridge the many academic works that we have surveyed, we consider a theoretical framework towards augmenting the ultimate display for rehabilitation. Such an augmentation would utilize the capabilities of a controlled iVE and quantifying emotion both through biofeedback (i.e., heart rate, sweat glands, and brainwaves) while also using in-game surveys to measure the user's self-perception and emotional state. The environment would factor in human perception and emotion through multiple co-dependent senses rather than a single sense. This could be achieved via olfactory modules, haptic feedback vests, and iVR HMDs. The system must account for pre-gameplay states and develop a baseline emotion profile for each user; this could be done by asking the user to relax for a set period of time while in the display in order to calibrate biofeedback sensors. With such a profile, we could examine how biofeedback changes occur when the user is presented with varying stimuli during exercise. The system may follow the effects of physical rehabilitation performance in comparison to biofeedback response and presented stimuli. By factoring in these metrics, we may be able to provide an iVR healthcare experience that adapts to each user's individual response and preferences.

This augmented display would equate to a sandbox controlled virtual environment to assist in the therapy process by enabling users to explore new attitudes, modulate cognitive biases, and examine behavioral responses. Through these multimodal sensory and motor simulations, researchers could craft experiences to assist in therapeutic engagement, and quantify or adapt the experience through biofeedback during runtime. Our vision for this augmented ultimate display comes from the synergy of three components: immersive media, biofeedback, and wearable robotics. Figure 2.1 demonstrates these mediums as inputs to augment the therapy process and show how they bring about emotional intelligence, physical intelligence, and adaptability.

As discussed in the previous sections, many components of this proposed augmentation have been rigorously researched independently within their respective fields. The synergies of these areas have the potential to produce emotional, physical, and adaptive intelligence from the interdisciplinary combination of these mediums. Nevertheless, these concepts are often not applied to healthcare. Some emergent research, as discussed in the previous sections (such as the work of i Badia's [168], has explored synergies between these areas, but these have not been fully demonstrated in healthcare or rehabilitation. Given the potential that immersive media has shown in therapy and rehabilitation, these fields and their synergies should be explored as one. This is necessary to advance the field of immersive media for healthcare and to fully understand how an ultimate display augmented for rehabilitation can be met. The center of Figure 2.1 represents this vision; a display in which the very world the user performs their rehabilitation in can adapt its difficulty and game mechanics to motivate and guide them through their emotional response through immersive computational media. Such a display would explore the limits of modeling a person's emotional reaction, mental perception, and physical ability, while also applying rehabilitation theory in a quantifiable and controlled environment. Just as the moon influences the tide, perhaps this display could influence our emotional "tides" to best perform rehabilitative tasks by influencing our perception for the better. The core elements of this biometric infused cyber-physical approach to immersive media in rehabilitation are illustrated in Figure

2.1.

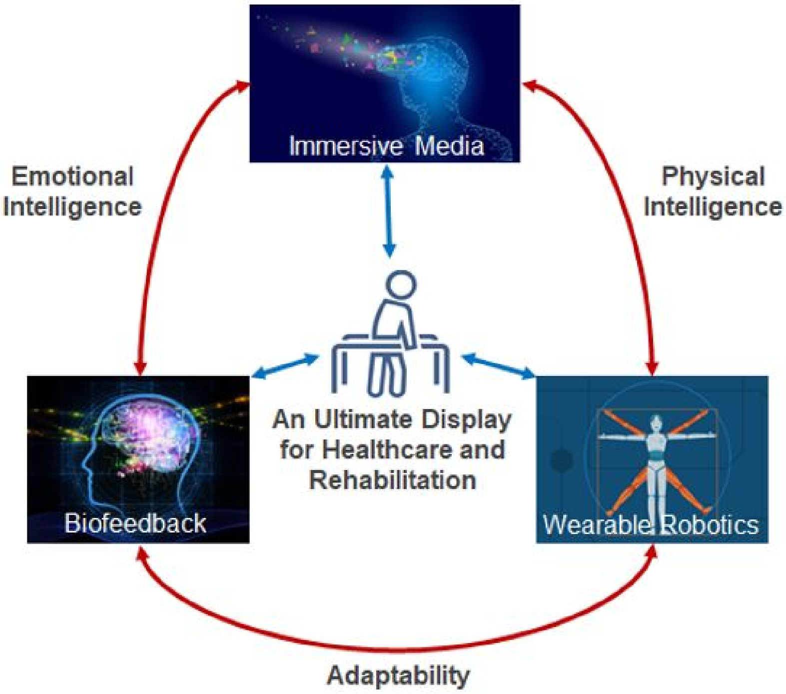


Figure 2.1: Components of the theoretical ultimate display augmented for rehabilitation. Areas and some of their synergies through immersive media , wearable robotics, and biofeedback. Elements of wearable robotics enable automated tracking of user progression. Physical intelligence examples include haptic stimulation, physical assistance, and positional sensing informed between the virtual experience and the wearable. Emotional intelligence examples include personalizing iVR stimuli by arousal response and calibrating the difficulty of iVR therapy based on heart rate. Adaptability examples include adjusting the physical assistance of wearable robotics and allowing for biometric input modalities to enable users of mixed ability to participate in the virtual experience.

This review examined how iVR can be a powerful tool in reducing discomfort and pain. As in the case of SnowWorld, created at University of Washington's HITLab, the experience demonstrated that iVR can be as effective as morphine in reducing pain for burn victims [61]. Much of this success can be attributed to the benefits and affordances of immersion [86, 87, 90, 69]. Therefore, the augmented ultimate display would need to enable the crafting of virtual worlds with high levels of presence and emotional engagement to assist user perception in overcoming adversity experienced in rehabilitation (such as pain and discomfort). One example to explore this may be readily feasible by augmenting the NIMH International Affective Databases (IAD) [8]. Researchers could extend these existing stimuli with multi-modality and evaluate user experience through biofeedback. Additionally, through utilizing the capabilities of a controlled iVE, emotion could be accurately quantified through both employing biofeedback while also using in-game surveys to measure the user's self-perception and emotional state. This data might be further explored to adapt both the immersive media stimuli and the level of assistance. For example, such an experience may allow researchers to build a baseline affective dataset for each user that could be applied to other immersive healthcare experiences with iVR. Similar emotional states from this baseline experience can be used to predict emotional response in order to adjust game difficulty and assist users with physical movement. Through this process, we may be able to create the ultimate behavioral sandbox for quantifying emotion during behavioral tasks and collect profiles to be applied to runtime physical therapy environments that can account for emotional intelligence during gameplay.

2.6.2 The ultimate display as a rehabilitation toolbox for task-based experiences

The development of an augmented ultimate display for rehabilitation may have broader impacts in the field of healthcare research. To illustrate some of the many theories that this system could explore, we share the following for consideration:

- Perception theory indicates that human perception is the composition of parallel senses of sight, hear, smell, feel, and taste, all of which influence behavior presence [176]. Subsequently, a multi-sensory iVR experience should induce more significant immersion with affordances for presence and emotional response [86, 87, 90, 69].

If this is true, perhaps we can create better iVR experiences for higher therapy engagement, compliance, and satisfaction.

- The Yerkes-Dodson Law states that, for any behavioral task, there is an optimal level of arousal to induce the optimal level of performance [139]. This law is one of the most frequently cited cognitive psychology theories but has never been verified [177]. If we can quantify arousal with the ultimate display by combining biofeedback sensing with in-game micro surveys, we may be able to verify the relationship between arousal and task-performance. If this is true, we may be able to create optimal stimuli to assist users in overcoming adversity within their therapy regimen.
- Csikszentmihalyi's Flow Theory suggests that total engagement in an activity can be achieved when perceived opportunities (challenges) are in balance with the action capabilities (skills) of an experience [178, 179]. This concept has been extended in virtual environments with "Gameflow," where user enjoyment is a result of balancing an environment's required concentration, challenge, skill, control, goals, feedback, immersion, and interaction of an environment [180]. Similarly to the Yerkes-Dodson Law, augmenting the ultimate display for physical rehabilitation enables a controlled environment to develop and measure optimal models of user engagement with therapy tasks.

2.6.3 Limitations of this review

There are many limitations to consider in this review. Firstly, the fields of rehabilitation, immersive media, and biofeedback are vast and ever-changing. However, we believe this review provides an adequate snapshot of the current potential that each of these literature review themes holds for assistive application. Additionally, this study primarily focused on iVR through head-mounted displays. Other extended reality mediums, such as spatial computing with augmented and mixed reality headsets, should be considered. With the advent of 5G edge computing and many extended reality devices exploring high-throughput streaming and social interaction, new paradigms for iVR-based therapy may emerge in the coming years. Yet, we believe that this review of iVR-based HMDs is still very relevant due to newfound consumer adoption and the necessity to drive and review the limitations of a field that is currently still maturing.

2.7 Conclusion

Immersive virtual reality paired with multi-modal stimuli and biofeedback for healthcare is an emerging field that is underexplored. Our bridging review of iVR contributes to the body of knowledge towards understanding immersive assistive technologies by reviewing the feasibility of a biometric-infused immersive media approach. We reviewed and discussed iVR therapy applications, the behavioral theory behind iVR, and quantification methods using biofeedback. Common limitations in all these fields include the need to develop a standard database for iVR-affective stimuli and the need for transparent dissemination of experimental methodology, tools, and user demographics in evaluating iVR for healthcare. We proposed an ultimate display augmented for rehabilitation that utilizes virtual reality by combining immersive media, biofeedback, and wearable robotics. Specific outcomes of such a system may include new algorithms and tools to integrate emotion feedback in iVR for researchers and therapists, discoveries of new relationships between emotion and action in physical therapy, and new methodologies to produce optimal therapy benefits for patients by incorporating immersive media and biometric feedback. These results may lead to deeper mediums for both clinical and at-home therapy. They may uncover novel approaches to rehabilitation and increase the affordability, accuracy, and accessibility of treatment. We believe that future of iVR healthcare may become a new field of therapy; a field that is centered on immersive physio-rehab that reacts, learns, and adapts its stimuli and difficulty to each individual user to establish a more engaging and impactful rehabilitation experience.

2.8 Acknowledgements

The authors thank Noah Wardrip-Fruin for his advices, recommendations, and support during this literature review. We also thank Richard Skarbez for his insightful peer review and helpful recommendations in the editing of this literature review.

Chapter 3

Accessibility needs of extended reality hardware: A mixed academic-industry reflection

3.1 Summary

The past five years have seen explosive growth in extended reality (XR) systems that make virtual, augmented, and mixed reality experiences available to the public. Emerging applications have begun to demonstrate XR's benefits from remote work, education, health, gaming, and so much more. But people with disabilities, who could benefit from these systems, are often an afterthought starting at the XR hardware level. In this experience report, we reflect on a year-long collaboration between various academic and industry organizations towards evaluating the current and future state of XR hardware accessibility. We draw on our group's experience of examining hardware as an inclusive team of XR professionals, both with and without disabilities. In doing so, we offer future directions for accessible XR hardware research and development.

3.2 Introduction

Extended Reality (XR) has redefined our relationships with the virtual world and each other – providing a means to augment how we work, play, heal, and engage by translating the physical world into the virtual [181]. Interaction within an immersive virtual environment has been found to yield a plethora of benefits for remote meetings, enterprise training, project management, entertainment such as tourism, well-being programs that stimulate mental and physical exercises, designing and testing products, and much more [181, 182, 72, 183, 184, 185, 186, 187]. For this paper, we emphasize a focus on XR Head-Mounted Displays (HMDs) systems that include the fields of Virtual Reality (VR), Augmented Reality (AR), and Mixed Reality (MR). The past decade has seen an explosive growth of these systems, with increases in computational power and affordability of digital systems effectively reducing barriers to technological manufacturing, consumer markets, required skills, and organizational needs [181, 186, 35]. These commercial systems provide a method for conveying 6-DoF information (position and rotation) to enable users to have a greater social-emotional interaction, which is especially advantageous for combating social isolation and assisting with remote work. However, XR's benefits through enhanced interaction modalities bring many challenges. These systems are challenged by the need for accessibility.

Since the advent of the internet itself, digital and online technologies have struggled with providing usage for people of all abilities [188]. An unfortunate truth is that most technologies are not designed for accessibility from the start. While some government policies are beginning to recognize and demand accessibility as a human right [189], access for all abilities is still often an afterthought with emerging technologies [190, 188]. The reality is that the XR industry is rapidly growing, has begun to define the virtual workplace, and is becoming a rich medium for social-emotional experiences. We must take steps to ensure usability such that XR technology can enable interaction for communities with disabilities. Much as how the internet has become an integral part of human life and work, the growth of XR may suggest a similar trend. Therefore there are many reasons to fight for XR accessibility: establishing equality of the digital and social workplace, engaging more users towards XR market growth, working towards more robust and user-friendly 3DUI interfaces, and enabling a greater user/developer diversity which is correlated to innovation and out-of-the-box thinking [191, 192].

To explore solutions for these challenges, we have spent the past year examining XR hardware accessibility needs with an interdisciplinary cross-industry group involving both people with and without disabilities through the XR Access Initiative. XR Access (XRA) is a community committed to making virtual, augmented, and mixed reality (XR) accessible to people with disabilities [193]. The goal for XRA is to engage with the academic and industry communities to create guidelines, and influence policy for accessibility in XR. Our group's focus is to support hardware, the starting point of most XR interaction capabilities, so that this technology is built from the ground up for all rather than application by application. In this paper, we reflect on our experience towards working to establish a future for accessible XR hardware. This consisted of an international workshop on XR hardware input modalities as well as a year's worth of monthly virtual working sessions. We discuss our findings from the debates, tensions, and questions we have had over the year in reflecting on the complexity and impact of creating accessible guidelines and developing calls to action. We aim to contribute our experience to offer considerations and future directions for impactful XR hardware research. In what follows, we briefly explore the state of XR hardware accessibility, contextualize our contributions, reflect on preliminary findings, and offer calls to action for future research.

3.3 Related Work

Many researchers have begun recognizing the need to examine XR for accessibility. For VR, authors have performed accessibility case studies to produce tools for people with low vision (alternate visuals settings [194]), high motion sensitivity (alternate navigation controls [195]), learning barriers (inclusive leadership training environment [196]), developmental disorders (museum storytelling [197]), and physical exercise (exercise games with low-cost exosuits [15]). Many authors have also performed accessibility studies to understand needs for AR: this includes wheelchair navigation for people with motor disabilities [198], image enhancement with text extraction for people with low vision [199], exercises for people with balance disabilities [200], and special education for individuals with cognitive disabilities [201, 202]. Additionally, some MR applications have explored the intersections of AR and VR with the physical world for accessible audio-visual navigation [203] and much more [204].

While much work has been done in exploring XR accessibility, much of these studies have been primarily software focused – addressing specific disability needs on an application to application basis. To establish a future with accessible XR experiences in mind, we argue that there is a need to approach XR accessibility at the hardware level. Hardware, after all, is what drives the capabilities of the software application from input methods, output modalities, control, and communication. By establishing an understanding of accessibility requirements at the hardware level, perhaps we can influence the default capabilities of XR software applications. Therefore, we reflect on what work has been done so far in XR hardware accessibility.

For the most part, there has been a plethora of research in examining custom hardware solutions and accessible controllers for XR experiences. These studies include configurable physical keyboards for VR [205], haptic and auditory canes to navigate complex virtual worlds without vision [206], bi-manual haptic actuators for greater perception of virtual objects [207], and hardware-software pipelines for rapid physical world mapping techniques to generate virtual environments [208]. Outside of XR, there has been an extensive amount of work in examining accessible hardware features for Automated Teller Machines (ATMs) [209], instructional technology [210], touch screens [211], wheelchairs [212], mobile devices in education [213], brain computer interfaces [214, 215, 216] and games [217]. While little work has been done in examining XR hardware accessibility, many industry and academic organizations are attempting to establish guidelines and recommendations for XR. This includes the World Wide Web Consortium (W3C) [218, 219, 220, 221], the XR Association [222], OpenXR [223], and Microsoft Research [224].

From our review of the existing published literature discussed in this section and experiences, we concluded that:

1. XR technologies have demonstrated a variety of benefits from industry to academia.
2. Current XR accessibility research has been primarily focused on the software level and many studies have demonstrated benefits for communities with disabilities.
3. Current XR hardware accessibility research has been primarily focused on alternate controllers and output modalities.
4. There is a lack of public standards for XR hardware accessibility and manufac-

turing.

5. Many of the published XR hardware studies require further validation and co-design with the disability community. It is critical to include communities with disabilities from the start of hardware design to ensure both usable and impactful XR systems [225].

With these challenges, we sought to form a team to engage with discussion, research, and networking to address these areas.

3.4 Contextualizing our Experiences

In July of 2019, the authors attended the XR Access Symposium in New York City, NY, USA [226, 193]. Over 140 people across academia, industry, and government got together at the Cornell Tech campus on Roosevelt Island for a day of presentations, demos, and breakout groups aimed at engaging XR Accessibility. Afterward, ten breakout groups were created to mix technologists, advocates, researchers, and industry leaders to brainstorm an initial set of goals that the XR Access initiative would consider. The mix of content, devices, and people enabled unique conversations that led to the next steps for XR Access. Our team was formed as an outcome of the input modalities breakout group during the XR Access 2019 Symposium [226].

We briefly introduce ourselves to help contextualize our experience and motivations with XR accessibility. Aviv Elor is an HCI doctoral student researcher that sparked his interests from a short-term disability after getting reconstructive surgery on his dominant arm as an undergraduate student. He engaged with VR to stimulate physical recovery during this period and became infatuated with exploring XR as a medium towards games for health [13, 12, 15, 227, 228]. Joel Ward is a technology strategist and innovator who has worked with accessibility since the early days of the World Wide Web and has been working with XR since 2016. This includes enabling and enhancing interactions for his son, who has disabilities, in the virtual and augmented worlds. Joel has also been speaking about the importance of accessibility for XR online, on his blog, inside his company, and "on the air" via the Workology podcast [229, 230].

Since the XR Access 2019 Symposium, the authors have formed and lead the XR Access Hardware Devices working group. This team expands upon the XR Accessibility efforts defined by the W3C [220], XR Association [222], and OpenXR [223] to evaluate the current and future state of XR hardware accessibility towards understanding guidelines and future needs. This consisted of monthly virtual workshop sessions from 24 contributing XR professionals with and without disabilities that representing the following affiliations (including the authors):

- Companies: Booz Allen Hamilton, Dell, Google, Hewlett-Packard, Magic Leap, Microsoft, Sony Interactive Entertainment, and Verizon.
- Organizations: AbleGamers, National Industries for the Blind, the Partnership on Employment & Accessible Technology (PEAT), and the XR Association.
- Universities: UC Santa Cruz and Cornell Tech.

All participants volunteered their time during these sessions to discuss challenges such as orienting XR hardware by accessibility needs, evaluating systems for accessibility, organizing involvement from XR hardware manufacturers & communities with disability, and brainstorming research programs to drive XRA hardware adoption. We should note that many more people were involved with the XRA Hardware Devices working group and are listed in the acknowledgments section. With our team contextualized, we discuss our working session organization and findings.

3.5 Working Group Design and Implementation

The XR Access Initiative community was organized as an outcome of the 2019 XRA symposium [226]. XRA brought together more than 140 advocates, researchers, technologists, and business leaders to advance the design, development, and production of accessible XR [193]. More specifically, working groups are organized by (1) Guidelines, Policies, & Practices, (2) Awareness & Outreach, (3) Education, (4) Application Accessibility, (5) Hardware Devices, and (6) Content & Authoring. Each group is overseen by two community leaders that organize monthly workshop sessions and meet bi-weekly with other group leaders through an executive team to facilitate collaboration.

The overarching mission of our initiative, as collaboratively defined with our meeting sessions, is to work with hardware manufacturers and academia to evaluate the current and future state of XR hardware while considering guidelines and user needs. To achieve this, we facilitated virtual monthly workshop meeting with our participants through video communication platforms such as Zoom (with appropriate accommodations upon request), and engage in discussion through communication platforms such as Slack, Twitter, Facebook, and LinkedIn. We additionally held collaborative meetings with the Guidelines & Application accessibility groups to establish synergies in addition to the executive meetings. All meeting recordings and working documents are publicly disseminated within the XR Access community through a shared google drive. In this next section, we reflect on our experiences from leading and guiding the XRA Hardware Devices working group during this past year.

3.6 Overview of the Preliminary Findings

At the start of our kickoff meeting, we reflected on the pilot work and hardware discussion from the XRA 2019 Symposium report [226] as well as our own experiences to define a set of working goals. Through our meeting and peer review with our sister groups, we established the following considerations for XR hardware accessibility research:

1. We must determine the state of XR Hardware Accessibility - evaluation of current devices, reaching out to synchronize with hardware companies current and future plans, working with academia to determine next-gen devices coming to fruition.

2. We must understand related fields of hardware, like general hardware accessibility, and what has been done before to ease accessible practices into XR hardware and expand upon the efforts of policy organizations such as the W3C and XR Association.

3. We must determine the most pressing problems in XR hardware accessibility - creating a community forum and hardware evaluation pipeline for involving communities with disabilities.

4. We must consider the Guidelines & Taxonomy for Accessible XR Hardware - working with industry and academia to create advice and a common terminology which shall also include functional performance criteria.

5. We must increase the public awareness of XR hardware accessibility - providing support resources, and incentives for XR hardware manufactures and developers alike.

3.6.1 Understanding the State of XR Hardware Accessibility

As described in our related works section, there has been a plethora of research that have begun to examine accessible design and inclusive tools for XR software accessibility. For example, in a study on "SeeingVR" by Zhao et al., the authors worked to produce a set of 14 tools that enhance an XR application for people with low vision and game developers by providing visual and audio augmentations through working with communities with disability [194]. The authors worked with both communities with disabilities and game engine developers for extensive evaluation and discussion of these tools with 11 participants and released an open-source game engine toolkit for developers to adopt these findings. In terms of hardware studies, research has primarily focused on alternative XR input/output methods and or adaptive sensing capabilities, as discussed in the related works such as custom vibrotactile walking canes to navigate virtual environments without vision. This research is exciting as it could potentially lead to breakthrough input methods such as Xbox's Adaptive Controller, which garnered adoption in communities with disabilities [231, 232]. However, little work has been done is establishing clear XR hardware usage requirements, and most studies still examine XR accessible on an application to an application basis.

Additionally, many policy organizations are taking note of XR accessibility needs and organizing workshops to establish directions for future research and guidelines. For example, the W3C has brought together over 170 participants to examine VR and the immersive web [219, 218, 220]. These participants from browser vendors, headset and hardware manufacturers, VR content providers, designers, and distributors contributed towards identifying standards for web-based XR systems and have begun to discuss hardware needs. The accessibility group considered input modalities and application APIs to assist with users with vision and hearing disabilities, while also providing extensive software recommendations in future workshops and WebXR resource standards. In the W3C 2016 workshop on WebVR, the executive report highlighted the need for VR hardware systems to "automatically adjust to meet the users restricted modality to bring controls closer if full range of motion is not possible" and "allow multiple ways to interact with the VR environment, and conversely multiple ways to extract information from it" [219]. In a position paper by Molt et al. of Microsoft Research, the authors examine the challenges of creating accessible VR and discuss the needs of content, interaction, input, interface, and device accessibility [224]. The authors argue on the need for inclusive representation of users within VR applications to both provide users with a means of increasing self-representation and control over how their avatars are presented to others. Aside from these discussions there was little to no talk on XR hardware accessibility. The breakout session argued that VR systems need to automatically adjust to meet user input modality and allow multiple ways to interact with the VR environments, but more work was needed into mapping these requirements for hardware. Given such, we have an understanding of how accessible XR hardware should act, but need to expand upon this work to establish generalizable XR hardware solutions and recommendations that are designed directly with the disability community.

3.6.2 Mapping Hardware Resources for Accessible XR Hardware Evaluation

From our discussion and collaboration with the guidelines & policy group, we have begun drafting evaluation metrics for physical hardware as well as operating system/platform input methods. This includes understanding the applicability of current XR HMDs for the following:

- Accessible Usability Needs: use without vision, limited vision, without color perception, without hearing, limited hearing, without speech, limited physical manipulation/strength, limited physical reach, photosensitive seizure triggers, and limited cognition.
- Hardware Components: operable parts, alternative input method controllers/hardware, biometric controls, status indicators, audio volume control, audio microphones, alphanumeric input keyboards, and keypads.
- Hardware Usage Types: simultaneous input, ease & force required to operate input, discernible tactile interfaces, tactile examination without activation, spacing, color, toggle of control status, visual label contrast & alternatives, and biometric ID/activation & input methods.

We should note that these considerations are a work in progress. Our group will actively continue to pursue our working goals, which may dynamically change and expand as our involvement expands to greater areas of communities with disabilities.

3.6.3 Questions, Debates, and Needs for Action

As we reflect on our community goals, mapped accessibility needs, and the state of XR hardware accessibility, we consider steps towards action. We have had many heated debates within our working group and would like to address some of them in consideration of future research. For starters: what came first, the chicken or the egg? Or rather, the chicken being manufacturing/evaluating accessible XR hardware and the egg being interest in accessible XR experience for communities with disability. At the current moment, the XR industry itself has quite a bit of maturing to do, and experiences by AbleGamers and similar organizations have often found a lack of interest in people with disabilities wanting to engage in co-design towards XR accessibility for some areas. We can develop guidelines for XR hardware accessibility, which can increase awareness, but these guidelines may not address the generative design thinking tools that developers need [217]. Additionally, we must consider the complexity and communicative nature of guidelines – we can establish a set of criteria to evaluate headsets. However, such guidelines cannot tell us whether that technology is accessible or impactful to the individual [217]. To paraphrase, if the Xbox Adaptive Controller is the most efficient tool to play a racing video game for someone with a disability, why suffer through the use of an XR platform that is inherently unusable for many disability needs?

So, where do we start? Does XR truly have value for the communities with disabilities? Moreover, is there any interest in accessible XR hardware? Our group has been engaging with these questions in considering our actions for the upcoming year. It boils down to a debate of universal design: do we consider XR hardware for the ”widest possible audience” (potentially limiting and severely restricting functionality) and or do we design XR hardware in such a way that usability can be modified, adaptable, dynamic, and flexible to user needs or custom hardware solutions [233]? To continue this line of thought, what sub-groups of communities with disability would be the most interested in engaging with XR hardware, and who would have the most gain in the short, medium, and long term? Lastly, if we are to establish evaluation methods and metrics of accessibility with XR hardware, how do we ensure that XR manufacturers adopt these concepts and are held accountable?

As discussed in the related works section, XR is slowly transforming our workplace, social life's, healthcare, and much more. Research has demonstrated benefits in many of these areas for communities with and without disability. As XR becomes mainstream, those who cannot use these devices may be left behind or required to use less immersive mediums at a disadvantage. In reflecting on these questions and the growth of XR hardware, we are organizing calls to action within the XR Access and the broader academic & industry communities.

3.6.4 Calls to Action

In this section, we highlight our calls to action for the research communities to work towards the future of XR accessibility. Based on our experiences and discussion of this paper, we present the following actions for consideration:

- We call for literature reviews and ethnographic studies centered around communities with disabilities in mapping XR accessibility and hardware research.
- We call for evaluating current XR systems to identify user needs, interests, and pressing problems for communities with disabilities.
- We call for the formation of a web-driven community directly connecting people with disabilities to XR developers in documenting and disseminating accessibility.

We are preparing to hold another workshop during the 2020 XR Access Symposium where we will discuss each of these areas and plan for action. Moreover, we are working with hardware manufactures and non-profits to organize a hardware loaner program to invite communities with disability to engage in XR hardware evaluation. We shall continue our cross-industry collaboration towards identifying active research projects and disseminating our findings in a public and accessible format. To this end, XR Access is expanding, and we invite involvement from everyone and hope to further diversify our team towards shaping the future of XR accessibility.

3.7 Conclusion

Extended Reality holds the potential to transform the virtual workplace, online social interaction, healthcare, and much more. However, people with disabilities are not able to access and benefit from these experiences as XR accessibility remains largely undefined and underdeveloped. In this paper, we have reflected on a year-long collaboration towards exploring the future of XR accessibility with an interdisciplinary cross-industry group of volunteers, both with and without disabilities. By engaging in a series of virtual working sessions and a symposium workshop, we offered a reflection of our experience in establishing goals, mapping out hardware, understanding accessibility, and calling action to all address XR accessibility. As we plan to engage in addressing these challenges collaboratively, we invite the academic community to engage in defining the future of XR accessibility. Accessibility is a human right – as we dive deeper into the immersive virtual world, we must ensure that people of all abilities can participate.

3.8 Acknowledgements

We acknowledge and thank the many active participants of the XR Access Hardware Devices working group. We additionally thank the XR Access Executive Team, especially Bill Curtis-Davidson, Dylan Fox, Devin Boyle, Shiri Azenkot, Larry Goldberg, Jessie Taft, and Liz Hyman.

Part II

CO-DESIGNING IMMERSIVE VIRTUAL REALITY GAMES FOR REHABILITATION EXERCISES

Chapter 1

Introduction

This part explores how immersive virtual reality head-mounted display systems can be utilized for serious game design in physical rehabilitation from [12, 13, 234]. We examine the iterative design and evaluation of two serious games that translate Constraint Induced Movement Therapy and Mirror Visual Feedback Therapy into exercise game mechanics for stakeholders with and without physical disabilities. Each case study concludes with design considerations for future researchers aiming to adopt immersive media for physical rehabilitation with head-mounted display systems, wearable robotics, and unique stakeholder needs.

Chapter 2

Project Star Catcher: A Novel Immersive Virtual Reality Experience for Upper Limb Rehabilitation

2.1 Summary

Modern immersive virtual reality experiences have the unique potential to motivate patients undergoing physical therapy for performing intensive repetitive task-based treatment and can be utilized to collect real-time user data to track adherence and compliance rates. This paper reports the design and evaluation of an immersive virtual reality game using the HTC Vive for upper limb rehabilitation, titled “Project Star Catcher” (PSC), aimed at users with hemiparesis. The game mechanics were adapted from modified Constraint Induced Therapy (mCIT), an established therapy method where users are asked to use the weaker arm by physically binding the stronger arm. Our adaptation changes the physical to psychological binding by providing various types of immersive stimulation to influence the use of the weaker arm. PSC was evaluated by

users with combined developmental and physical impairments as well as stroke survivors. The results suggest that we were successful in providing a motivating experience for performing mCIT as well as a cost-effective solution for real-time data capture during therapy. We conclude the paper with a set of considerations for immersive virtual reality therapy game design.

2.2 Introduction

Long-term intensive and expensive physical therapy is often necessary for patients with orthopedic, neuromuscular, cardiovascular, pediatric, oncological, and other conditions. There is significant need for cost-effective yet efficacious treatments to address such conditions [235]. In 2014, Medicare spent \$7 billions on inpatient rehabilitation therapy (which includes physical therapy) for 339,000 patients. Currently, Medicare caps outpatient spending at only \$1900 per person per year, which can be spent on physical therapy as well as speech and language pathology payments. This does not cover the needs of all patients; 19% of beneficiaries with complex impairments exceed Medicare's cap, spending on average \$3,013 out-of-pocket [236]. Such costs can be reduced by moving inpatient treatment to outpatient programs, which include exercises performed independently by patients outside of the therapy office.

The main issue with relying on patient-initiated exercises outside of the therapy office, however, is that a patient's progress is often dependent on their ability to adhere and comply with such exercise program. The compliance rate for such programs can be quite low, due to perceived barriers to exercise, lack of positive feedback, and degree of helplessness [40, 41]. Patients who set goals collaboratively with their physical therapist have higher compliance rates [237]. For the purpose of our study, the operational definitions for adherence is a patient's ability to perform exercises as prescribed to them, and compliance is the ability for a patient to continue exercising for the frequency and duration of their prescribed rehabilitation program.

While home regimens traditionally involve a physical therapist's verbal instructions and printed pictures, higher-tech initiatives associated with telemedicine, virtual reality and robotic programs have been shown to be more effective in promoting compliance than traditional paper-based and verbal instructions [45, 46, 47]. These higher-tech exercise programs often use sensors, either to passively monitor a patient's status, provide feedback so an action can be modified, or use actuators to assist the patient in completing a motion [48, 49]. The sensors can include accelerometers, gyroscopes, microphones, bio-chemical sensors, motion detectors, GPS, and EMG. Several Kinect-based systems were proposed for rehabilitation, using avatars to portray the correct motion, or using therapy-based games, e.g. [238]. The feedback provided by these systems included auditory, such as music or beeps; tactile, such as vibrational cues; and visual, with a screen showing correct position or pressure color indicators. While these systems are sophisticated, our literature review suggests that they are not optimal for rehabilitation because they were designed for entertainment instead of specifically targeting patients performing prescribed health intervention practices.

The system reported in this paper is an exercise game for upper limb rehabilitation, which can be set up inside or outside the therapy office by a caregiver or family member, with the following characteristics:

1. It employs an immersive Virtual Reality (VR) game as a motivation-inducing medium.

2. It provides immediate feedback to users about their exercise performance through

real-time motion tracking.

3. Its exercise program is adapted from an established physical therapy protocol called “modified Constraint Induced Therapy” (mCIT) [239].

Through our literature search, we believe we are one of the first to leverage the HTC Vive, a recently available commercial VR system, as a single system to integrate high resolution real-time data tracking and collection with a therapist interface in the specific application of upper limb rehabilitation for hemiparesis (weakness in one arm). We are also one of the first to use immersive Virtual Reality to provide psychological motivation to increase compliance in an established method of mCIT with a relatively large sample of users (for example, the closest work that uses Virtual Reality for CIT had participation from only four stroke survivors [240], whereas the study discussed in this paper includes participation from nine stroke survivors as well as six users with developmental disabilities with upper-limb impairments). The therapeutic intent of our study is to offer a version of mCIT that is more engaging, accessible, and uses recorded real time tracking, which could result in improved user experience and increased efficacy in rehabilitation improvements for a given time.

2.3 Why Virtual Reality?

The term ‘Virtual Reality’ was coined long before the advent of recent immersive virtual reality (iVR) systems, which brought differences in opinions of the definitions of the terms ‘’immersion“ and ‘’virtual reality“. For this study, immersion reflects how well the user is connected to the virtual environment which is dependent on the number of motor channels and range of inputs provided to achieve a high enough fidelity of sensory stimulation [88]. VR systems that provide a head mounted display (HMD) with a binocular omni-orientation monitor along with appropriate three-dimensional spatialized sound were categorized as iVR systems.

Studies comparing VR and non-VR systems for rehabilitation had largely shown that the VR version outperforms the non-VR comparable. A study of upper limb rehabilitation in patients with hemiparesis had shown that the VR group displayed significantly improved performance in paretic arm speed [241]. In another study of using 2D video-capture virtual reality (VR) training environment to improve upper limb motor ability in stroke patients compared to those performed in conventional therapy, more patients in the VR group improved upper limb clinical impairment and activity scores and the improvements also occurred earlier [242].

Therapeutic VR applications using iVR, non-immersive virtual reality (niVR), and augmented (mixed) reality (AR). niVR systems, or desktop systems, utilize a monitor and allow interaction of the user through conventional means such as keyboards, mice, or controllers [33]. AR systems employ virtual feedback by allowing the user to see themselves and their surroundings projected virtually onto a screen, usually in a mirror like fashion [54]. These systems are similar in how they present movement based tasks with supplementary visual and auditory feedback, but differ in their interaction methods [55].

There has been a large number of publications and studies focusing on VR technologies and the applications in psychotherapy, physiotherapy, and telerehabilitation within the past two decades. In prior literature, VR has been defined as a computer user-interface that uses real-time simulations to invoke user interaction [72]. Modern VR technology is commonly known for its impact on enhancing the video gaming paradigm through involving more dedicated user interaction [56]. The increased physical demands of these video gaming platforms has garnered interest for their potential in therapy through repetitive and quantifiable motor learning protocols [57]. The use of VR systems has been shown to be effective for emotional, psychological, and physical therapy [58, 59]. Environments can be tailored to cue specific movements in real-time through sensory feedback via the vestibular system and mirror imagery to exemplify desired ranges of motion [73].

2.3.1 Virtual Reality Applications For Psychological Therapy

Psychological research has seen an increasing use of VR due to its ability to simulate realistic and complex situations critical to laboratory based human behavior investigations [69].

Some of these investigations include the success in reducing pain through the use of stimuli in VR that is equivalent to effects of an analgesic [61, 62]. With the immersive capabilities of modern headsets such as the HTC Vive and Oculus Rift, there has been an increase of success for VR exposure therapies such as for post traumatic stress disorder [63, 64], borderline personality disorder [65], phobias [66, 67, 69], and schizophrenia [68] as well as many other psychological therapies.

The acceleration of VR use in psychological therapy can be attributed to the relationship between increased presence and emotion [64]. Increasing the amount of stimuli using iVR is a key factor to influencing user behavior and experience [70], and, with the price of computing devices and hardware decreasing, VR headsets are becoming more popular and immersive [243, 244]. Thus increasing immersion through VR leads to more emotional influence of the user and can incite desired physiological responses through crafting a stimulating and engaging virtual environment [71].

2.3.2 Virtual Reality Applications For Physical Therapy

Traditional forms of physical therapy and rehabilitation are based on therapist observation and judgment, coincidentally this process can be inaccurate, expensive, and non-timely [42]. There is evidence that VR can be an effective tool in improving outcomes compared to conventional physical therapy [72]. With the emergence of new immersible multimedia, VR experiences with sight, sound, and touch can be augmented in rehabilitation. As a result, VR has been shown to be effective in improving a variety of motor impairments, such as hemiparesis, caused by Parkinson‘s disease, multiple sclerosis, cerebral palsy, and stroke [73].

High repetition of task oriented exercises are critical for locomotive recovery, and user adherence to therapy protocol is imperative. VR based physical rehabilitation can induce adherence to therapy protocol as well as and sometimes better than human-supervised protocol due to the capabilities of multi-sensory real time feedback [50]. The multi-sensory, auditory and visual, feedback can further persuade users to exercise harder through increased stimuli [50].

VR based physical rehabilitation also allows for increased quantitative feedback for both the user and the therapist. With modern iVR systems implementing three dimensional motion tracking, VR serves as an effective way to monitor progress during rehabilitation. [56].

2.3.3 Teletherapy Virtual Reality Applications

Telerehabilitation can result in decreased treatment cost, increased access for patients, and increased quantifiable data for therapists [79]. There have been various studies confirming the technical feasibility and efficiency of in-home telerehabilitation with VR [80]. Users generally achieve better results in rehabilitation due to the increased feedback from the telerehabilitation VR [81]. Telerehabilitation VR increases the usability and motivation of the user compared to traditional in-home therapy systems, while reducing work for physical therapists and reducing costs for patients [82].

2.3.4 Virtual Reality and Motor Stroke Rehabilitation

Our literature review concludes that VR is largely effective for motor stroke rehabilitation [42, 74, 75], with many of these studies supporting that the use of VR had resulted in significant improvements when compared to traditional forms of motor stroke therapy [73, 50]. These studies used Kinect, Nintendo Wii, IREX: Immersive Rehabilitation Exercise, Playstation EyeToy, Cave Automated Virtual Environment, as well as custom designed systems. For a given treatment time, video game-based rehabilitation is more effective than standard rehabilitation [50, 73, 72]. The physical rehabilitation communities have been enthusiastic with the potential to use gaming to motivate post-stroke individuals to perform intensive repetitive task-based therapy, with some combining motion capture as a way to track therapy adherence and progress. However, older studies commented that technology at the time needed to improve, such as in motion tracking accuracy, to become more effective, reliable, and accessible [77, 42]. Additionally, more studies are needed to continue gaining a deeper understanding of the efficacy of VR in rehabilitation [74, 78].

Interactive video games and VR games were applied in stroke rehabilitation as early as the era of the Sony PlayStation II gaming console, with the EyeToy Camera for motion capture [245], and the Nintendo Wii [246]. The combination of games and motion capture technology was used successfully in home and clinical settings, demonstrating the feasibility and potential of such technology. Since then, a variety of games employing robots, tangible surfaces, and VR systems have been proposed for stroke rehabilitation. Specific applications include a robotic-assisted game for motor and cognitive post-stroke rehabilitation [247], a multi-user VR game for upper extremity rehabilitation [248], and a tangible gaming board for the training of upper limb gross and fine motor skills [249].

However, applying consoles and games, which were designed for entertainment applications, to stroke rehabilitation has several challenges. The first challenge is the design of the controller, which might not work for individuals who have insufficient hand and grip strength or arm control. The second challenge is understanding how the controller works to tailor the game mechanics appropriately for a particular therapy goal. For example, the Wii remote responds to acceleration more than position changes, so faster movements will garner more points than wide-sweeping but slower movements [250]. This mechanic might not translate well for stroke rehabilitation that encourages users to extend their arms as much as they can as they are not rewarded for having wider ranges of motion. The third challenge is for the game itself to strike many balances, including its:

1. Levels of difficulty: the game should be adjustable for different ability levels that can vary depending on the severity of hemiparesis.
2. Rules/mechanics: the game should not assume any prior knowledge of rules and should be understandable by those who might be cognitively impaired, which is a

common symptom of stroke affected persons.

3. Design: the game's theme, reward, and design elements such as background and soundtrack should not be perceived as too childish, boring, or meaningless by users of varying ages and backgrounds.

4. Data interpretation: Data should be leveraged to measure progress towards therapy goals and be used for treatment planning.

Past motor rehabilitation studies have been far and apart due to costly and sluggish VR systems existing in the past [33]. These issues are being resolved, leading to high performance VR headsets costing only a fraction of the price; with new headsets such as the Oculus Rift, PlayStation VR, HTC Vive, Google Cardboard, and Microsoft Hololens [243].

2.4 Why Constraint Induced Movement Therapy?

As mentioned above, the purpose of the game we develop is to act as a motivating exercise program for upper limb rehabilitation. To inform the game mechanics, we leverage an established therapy technique commonly used for upper limb rehabilitation: constraint induced movement therapy.

Constraint induced movement therapy (CIMT) involves massed and intensive practice with the weaker affected upper extremity and includes two components: the use of the unaffected upper extremity is restrained during 90% of waking hours, and at the same time, the more affected upper extremity receives repeated and intensive training for more than six hours per day [239]. CIMT has been widely used and studied compared to traditional rehabilitation techniques and "could improve functional performance and increase the usage of the more affected upper extremity" [239]. Although research shows benefits from CIMT, in a survey of stroke survivors, 68% of respondents said they were unlikely to comply with the therapy protocol due to either logistical aspects (length and duration of therapy) or aspects of the therapy itself (wearing a physical constraint for a long period) [251].

Modified constraint-induced therapy (mCIT) is a form of CIMT that requires less engagement and compliance from the patient. Researchers have designed a modified CIMT that has a shorter intensive training period as well as shortening of the period that the unaffected upper extremity is constrained [252]. For example, a patient may visit a therapist several times per week and in each thirty minute session the patient practices focused exercises using their weak arm. This therapy has been demonstrated to increase the mobility and use of the patient's arm only if they have some mobility remaining in their wrist and fingers [252]. The program our system was designed for follows mCIT instead of CIMT, with a slight modification. Instead of a physical constraint, we utilize the game mechanics to provide a psychological constraint. Specifically, we provide more reward if the participant uses their weak arm, and considerably less reward if the participant uses their strong arm during game play.

2.5 Our Study

Our study aims to evaluate the use VR for rehabilitation, specifically the use of the HTC Vive for upper limb rehabilitation of users with hemiparesis. To inform our study, several commercial, private, and research based systems have been methodically tested, including Microsoft Hololens, Microsoft Kinect, Oculus Rift, Playstation VR and HTC Vive. We came to the conclusion that HTC Vive provides the most high resolution data capture regarding head and upper limb movements. It also provides the most accurate 6-DOF motion capture capabilities [253]. More detailed explanation of why HTC Vive was the chosen technology is described in the System Design section.

The aim of our study can be translated into four main questions:

1. Are participants with hemiparesis able to use the controllers during game play?
2. Can the game mechanics result in improved adherence and compliance than the adherence of conventional mCIT as reported in the papers?
3. Is the game understandable and enjoyable by our target population?
4. What does the data tell us about user game play strategies?

We sought to answer those questions through a series of brainstorming sessions with physical therapists, stroke survivors, their caregivers, and us at Cabrillo College‘s Stroke and Disability Learning Center (SDLC) and Hope Services, our long-term collaborators. The intent of mCIT is to constrain the stronger limb in order to encourage the use of a weaker limb. Our approach makes use of a VR game with the same intent. Therapists concluded that the benefits outweigh the risk of using a psychological constraint in VR instead of a physical constraint. While they didn't expect a VR game to cause injury to patients because game play can be done seated and physical interaction with the game is limited to a lightweight headset and controllers, the therapists said there is risk that the VR rehabilitation game potentially won't be effective in encouraging use of the weaker limb and may not be enjoyed by patients, especially if they have never experienced VR before. Because our VR game is dependent on psychologically encouraging the use of a weaker limb, it could be rendered ineffective if the gaming experience does not motivate patients. To mitigate this risk, we considered game themes in unconventional environments that could distract or fascinate the player. The benefits of the VR rehabilitation game is the prospect of using the game at home, which allows patients to maintain treatment even at times when they can't make it to clinic sessions, and that psychological constraint doesn't feel as invasive to the independence of the patient as a physical one does. Physical constraint, therapists commented, could feel demeaning and makes some patients feel like they lose independence, so they may resist treatment especially in the beginning before they start to feel the benefits of mCIT. Patients must go to a clinic and are dependent on being strapped into their restraint by a therapist, and the restraint itself could feel claustrophobic or unconformable. The therapists believed that psychological constraint through game rules and an immersive experience wouldn‘t feel like losing independence, instead, the patients maintain dignity, and feel as if they chose to play the game. Patients can even measure their own progress by tracking game scores and session durations. Thus we expected that the use of psychological constraint for a gamified mCIT protocol would be equivalent or more successful than using physical constraints.

We went through the process of choosing the game domain that appeals most patients. We defined the following characteristics for our game:

- The game domain chosen is about stars and galaxies.
- The game provides immediate feedback about users‘ task performances through data captured using motion tracking sensors,which can allow patients and therapists to see performance in real time.
- The game encourages the use of hemiparesis-affected limbs through adapting the principles of an established physical therapy protocol called “modified Constraint

Induced Therapy" (mCIT) described above [239].

We should note that the long-term goal of physical therapy is to gain back the ranges of motion and strength of the weak arm [254] while the scope of this study is to encourage users to use their weak arms while capturing movement data that can be used by therapist. The Hope Services and SDLC therapists reviewed and approved our proposed game play mechanics and worked with us to develop the user evaluation protocol.

2.6 System Design

Our designed game is named, for simplicity, Project Star Catcher (PSC). PSC incentivizes game participants to use their hemiparesis-affected arm through catching falling stars in a cosmic VR environment while collecting real-time user data. The game was developed leveraging the HTC Vive system that includes of a headset, wireless controllers, and lighthouses.

Upon reviewing multiple commercial VR and motion capture systems such as the Oculus Rift, Playstation Morpheus, IREX and Optitrack, we decided to use the HTC Vive as the medium to create a upper limb rehabilitative experience. The HTC Vive differs from other modern systems in its localization system and affordability. The HTC Vive, developed by the Valve Corporation and HTC, is a VR head mounted display (HMD) that implements room scale 4x4 m tracking technology by utilizing a "lighthouse" system of lasers which enables user interaction and accurate 3D motion capture (+/- 1.2mm worst case) for an affordable price of 600 USD [255]. Its adjustable headset allows for users who wear glasses, which might be the case of many of our target population [256]. Complementary to the HTC Vive are two hand held controllers that feature dynamic haptic feedback which enhance spatial orientation [257]. The HMD provides a 110° field of view and 90 Hz refresh rate, and the system itself has a significantly low latency of 22 milliseconds, which is known to reduce motion sickness [258]. Motion capture is tracked at 120 Hz using infrared laser sweeping and photodiodes that enable for recovery of position and orientation [257]. The HTC Vive also features a safety guidance system preventing users from potential injury in the real world environment [258, 259]. While the HTC Vive is known to produce inaccurate data and reset when the user exits the tracking area [259], the risk of this affecting our therapy game is minimal. Because our target participants are persons with motor impairments that need upper-limb exercises, we can prescribe VR interactions to be in a seating position and contained within the tracking area.

In PSC game play, the HTC Vive controllers are visualized in the player‘s hands as “Star Catchers” that are colored either red (for the strong arm) or green (for the weak arm), seen in Figure 2.1. The Star Catcher colors were chosen to help remind users that their goal is to catch with the weak arm. Before the player is an ethereal, cosmic galaxy that the stars fall from. Stars are caught when the player touches them with a Star Catcher, resulting in haptic feedback as the controller vibrates, figuratively feeling the stars being caught, and an attached speaker plays sounds effects. This design supports an immersive experience of sight, sounds, and touch that could be exciting enough to motivate users.

PSC has several levels of difficulties in terms of the speed of the falling stars, the rate of the falling stars, and the angles the stars fall from. For a personalized experience, the difficulty of the game can be adjusted by the evaluator or user based on the data collected in an introductory round, the user‘s past rehabilitation history,

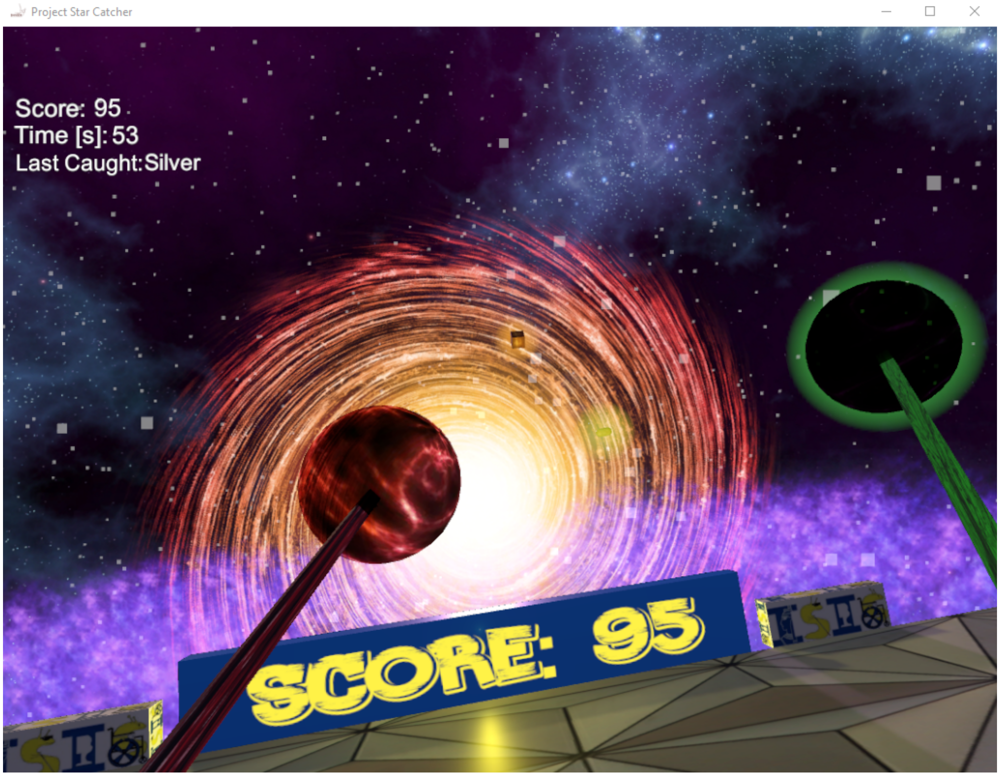
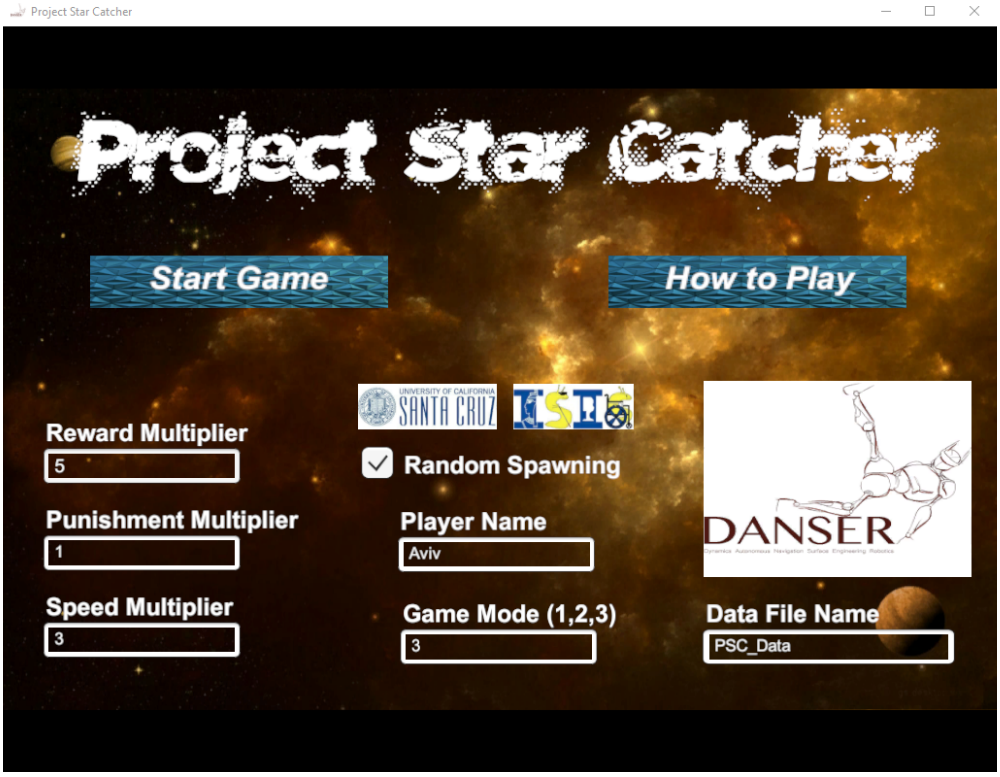


Figure 2.1: Player Gazes into the Cosmic Galaxy Showing Red and Green Star Catchers as Arms Lift to Catch Stars

or the recommendation of a therapist. The stars are colored bronze, silver and gold to indicate different speeds of falling and score values. The base speed is adapted to each individual player‘s reaction time. Bronze star falls at 75% of the base speed, silver at 100%, and gold at 125%. The score is calculated based upon the user‘s performance, with the weak arm (green Star Catcher) providing the most weight, and it is mainly used as a motivational landmark for the user to see progress and beat high scores. Stars spawn from various locations to make sure that users will perform rehabilitative arm movements to catch them. These movements were designed in consultation with a physical therapists from the SDLC and Hope Services.

PSC was developed in Unity v5.5.0f3 with the SteamVR plugin v1.2.0, which connects HTC Vive hardware to Unity. SteamVR is open source and both have documentation for C# and Javascript. Using the Unity physics engine, the Rigidbody class was used to model stars and the HTC Vive controllers were modeled as Capsule Colliders that detected contact with stars. Stars randomly spawn at different locations along a half circle in front of the player in three game modes - stars fall from 0^o (directly above the user) in Mode 1, 45^o in Mode 2, 90^o (directly in front of the user) in Mode 3 - to vary motions and body posture when catching stars. All game play was set to a time-dependent state, allowing for a customizable and time-dependent assignment of the moving stars. Data communication was achieved through utilizing C# and Microsoft .NET Framework allowing for successful data exportation of 90Hz (and greater depending on the machine).

The evaluator user interface, shown in Figure 2.2 has the following inputs: Reward Multiplier, Punishment Multiplier, Speed Multiplier, Game Mode, Player Name, and Data File Name. Reward and Punishment Multipliers allow a therapist to cus-



Reward Multiplier: 5

Punishment Multiplier: 1

Speed Multiplier: 3

☑ Random Spawning

Player Name: Aviv

Game Mode (1,2,3): 3

Data File Name: PSC_Data

Figure 2.2: Project Star Catcher Evaluator Interface (note that the "How to Play" module contains explanation about what the game settings are and how to adjust them)

tomize the amount of score points earned for the green and red Star Catchers, respectively. Game Mode allows selecting Mode 1, Mode 2, or Mode 3. Player Name allows PSC to provide identifying information to data collected during game play.

These mechanics were implemented so that the user could perform therapy without a therapist and occasionally meet a therapist virtually so that they could be re-evaluated, effectively enabling low-cost telerehabilitation through readily available online collaboration platforms.

2.7 Pilot Study

To investigate the evaluation protocol and ensure that our study and game design can answer the four questions we set out, we evaluated PSC with six participants with developmental disabilities who have upper limb impairment. The participants consisted of four men and two women, with ages ranging from 23 to 31 as seen in Table 2.1. We received IRB approval from our university for our game play testing.

Participants were recruited through Hope Services, a local non-profit organization that promotes independence for people with a variety of developmental disabilities, including Autistic Spectrum Disorder, Down Syndrome, and cerebral palsy. By providing jobs and other activities, Hope Services intends to change attitudes about the abilities of these persons in the workforce and in the community. The session started with the participants filling in a consent form as well as a video and audio release form. A professional caregiver was present to help explain in a way that is understandable to the participants and make sure that the participants did not get overstimulated or did any movements that might cause upper limb strain. There were always two to three researchers present in the sessions. The researchers and caregiver helped the participants don the head-mounted display and place the Star Catchers in their hands if requested. However, most participants were able to don the head-mounted display by themselves.

After getting an explanation of game, the players were introduced to Modes 1, 2, and 3 for about 1 minute each. The players then selected their preferred mode and played for 5 minutes, a duration recommended by therapists at SDLC and Hope Services. During the 1 minute introductory modes, the participants' reaction times were calculated and a base speed for the falling stars were customized to each user for their 5 minute trial so that the stars were not too slow or too fast.

Table 2.1: Summary of Participant Demographics Pilot Study

ID	Age	Gender	Disability
Co	23	M	Cerebral Palsy
Ty	31	M	Autism
Pp	25	M	Inverted X Syndrome
Cn	24	M	Autism
Cs	26	M	Mild Intellectual Disability
Fa	30	F	Down Syndrome

The play session was videotaped. After the session, the participants were interviewed using a retrospective testing method, in which the researcher and the participant reviewed the video together, and the reviewer stopped the video at multiple points and asked the participant directly to explain certain behaviors when deemed necessary.

2.8 Results and Discussion from the Pilot Study

The pilot study aims at investigating whether our game design and study protocol can answer four main questions:

1. Are participants with hemiparesis able to use the controllers during game play?
2. Can the game mechanics result in improved adherence and compliance than that of conventional mCIT?
3. Is the game understandable and enjoyable by our target population?
4. What does the data tell us about user game play strategies?

To answer these questions, we analyzed the collected performance and behavior data.

Figure 2.3: An example of a participant's initial strategy in holding the controllers

2.8.1 Are participants with hemiparesis able to use the controllers during game play?

Our concern in this study was that whether the participants could hold or manipulate the controllers.

Direct observation of the sessions revealed that the six participants were able to hold or manipulate the controllers enough to play PSC. Depending on their disability, they developed user-specific strategies to hold the controllers better. Figure 2.3 shows one participant with cerebral palsy that has weakness and hypertonicity in his right hand, who held the right controller by resting his hypertonic pointing finger inside the ring.

There was one occasion when the controller slipped from a participants' hand and fell to the ground because this participant with wheelchair was very excited and made an exaggerated movement that caused the controller to fly. We found that wrist straps prevented this incidence and hence, since that point, we ensured that participants had wrist straps put on before starting the game play.

We also deduced that arm sleeves can also help the controllers stay on their hands (arm sleeves covered from midway between their elbows and hands all the way to their palms, with the controllers' sticks encased inside those sleeves).

2.8.2 Can the game mechanics result in improved adherence and compliance than those of conventional mCIT reported in the papers?

PSC was designed to improve adherence and compliance rate in a psychological constraint induced therapy by using game mechanics to provide reward if the participants use their weak arms to catch falling stars of varying speeds. It was apparent that we could easily gauge adherence rate through analyzing performance data that the motion capture system that came with HTC Vive collected. As for compliance, a post-testing survey showed that 84% of participants believed they were 'very likely to comply' with using a VR game like PSC as a long term rehabilitation tool.

To do a proper comparison of the adherence rate, we normalized the number of the stars caught by each participant into percentages for a specific star color. So for example, if a participant caught 20 bronze stars with the weak arm and 20 with the strong arm in the five-minute session, then it's calculated as 50-50.

Figure 2.4 shows the percentages of the stars caught by the six participants, each using the weak (green) and strong (red) arms while playing the first iteration of

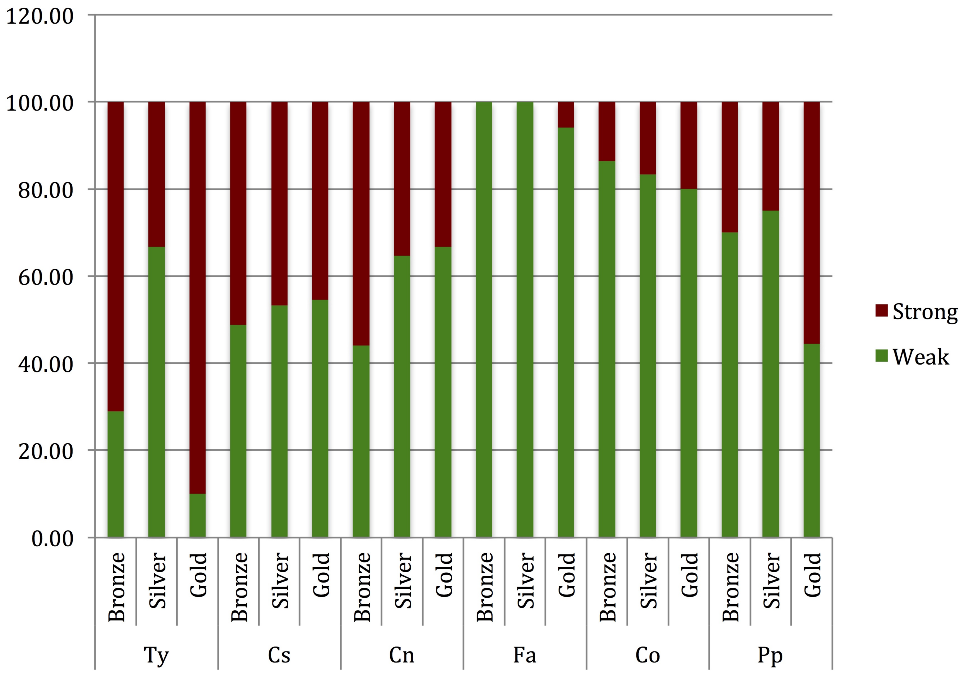


Figure 2.4: The Percentages of Stars Caught by Each Participant in Pilot Study

PSC. In general, we can see that the participants are mostly adherent to the game rules. Specifically, they tried to use the weak arms as much as they can. The exception is Participant Ty, who slipped into using mostly the strong arm for the gold stars (that are the fastest of the star types). Participant Fa almost had the perfect compliance, having only one gold star caught with the strong arm, and Participant Co was also highly compliant, although again, just like Participant Ty, slipped into using the strong arm in gold star condition.

Figure 2.5 shows the median of the percentages of these six participants by star color while playing PSC. This figure overall shows that participants are adherent to using their weak arms, averaging 60%. This percentage is considerably higher than the reported adherence rate of conventional mCIT, where 32% of patients were likely

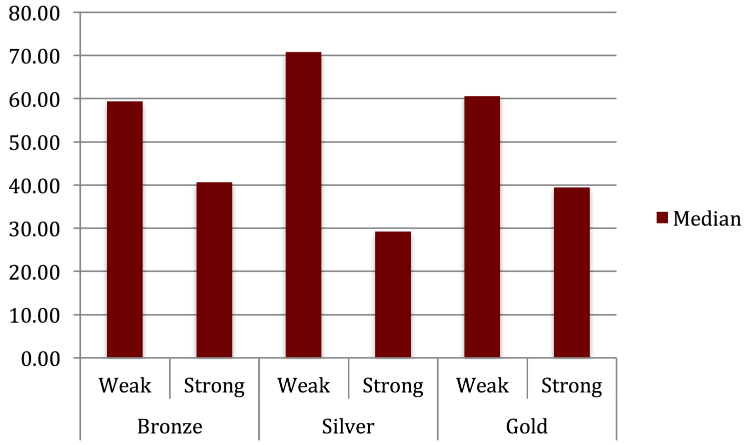


Figure 2.5: Median Percentages of the Stars Caught in Pilot Study

to use traditional therapy protocol [251].

2.8.3 Is the game understandable and enjoyable by our target population?

To answer this question, we relied on qualitative data. The qualitative data reported in this study consists of a combination of (i) themes extracted from the interviews and (ii) observations obtained from the researchers during the play sessions. The interviews were transcribed and analyzed by two researchers using content analysis, more specifically open coding, to extract themes. The researchers did the open coding in a line-by-line fashion. The researchers came into the data with no preset theme or bias. The observations were extracted by two researchers watching the videos together but coded the behaviors independently. The two video coders then compared the behaviors and resolved differences in coding through discussing the cause of the difference to come up with unified themes.

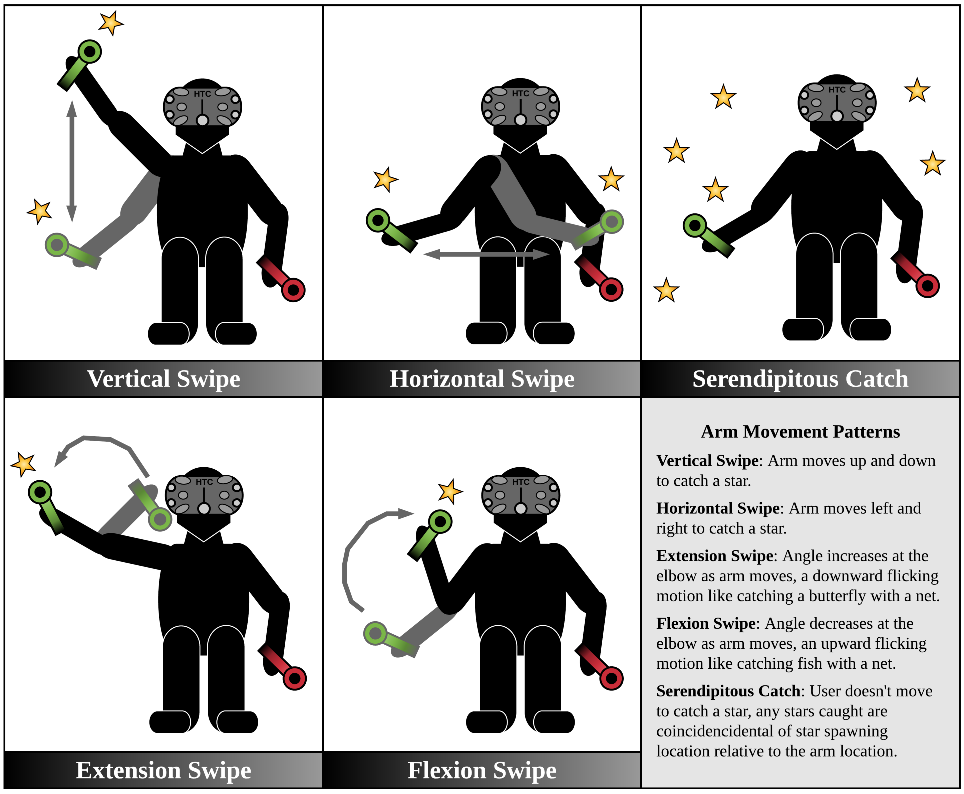
During game play, we noted that the HTC Vive did not cause noticeable tracking jitter and participants confirmed that they didn't feel motion sickness. Participants thought PSC was fun and enjoyable, and did not request significant changes or improvements to PSC other than "happier music." They did request "more games" with different themes because it would be "more interesting." When asked what they thought of the sound effects and haptic feedback on the controller, most participants stated they did not even notice the music or vibration. Similarly, when asked whether they felt that the star colors were visible, most did not notice that there were three different stars, although they noticed that some stars fell faster than other stars. Finally, when asked whether they preferred the game or the physical therapy exercises that they currently do, everyone unanimously preferred PSC, which they concluded was more fun and entertaining than their usual exercises.

2.8.4 What does the data tell us about user game play strategies?

To answer whether the participants understood the game mechanics, we relied on observations from game play that was recorded and analyzed post-session. We captured game play data automatically by counting the number of stars caught with weak and strong arms. Our video data was analyzed to derive strategies and performance. Video coding a 5 minute session video took two coders an average of 20 minutes. Transcribing and analyzing the interviews took another 20 minutes average per participant per coder. Through game play data video content analysis, we extracted several themes about these strategies. The observations yielded a rich set of results that indicated that not only participants understood the game mechanics, they developed strategies to more effectively catch the stars in spite of their upper limb weaknesses.

Through video content analysis, we extracted several themes about these strategies:

1. Number of strategies employed = [1, 2-3, 3+]. Some participants were consistent in their strategy in catching the stars while others changed strategies multiple times in the span of 2 minutes. For example, one participant changed strategies more than 10 times.
2. Arm ranges of motions = [only lower arms (resting the elbows somewhere), the whole arms]. Essentially, some participants rested the elbows on their wheelchairs or chair arms, only moving the lower arms to catch the stars, while others swung freely their whole arms .
3. Arm movement patterns = [vertical swipe, horizontal swipe,extension swipe, flexion swipe, serendipitous catch]. The basic motions are illustrated in Figure 2.6.
4. Arm movement speeds = [fast, slow]. We noticed that participants either waited by resting their arms and then extended one arm very fast to catch the star [fast] or slowly extended one arm anticipating the location the star would hit the Catcher [slow].
5. Head movements = [vertical rotations (e.g., nodding), horizontal rotations (e.g., shaking heads), looking only at a certain angle (e.g., either up, down, or straight ahead)]. For example, one participant in wheelchair rested his head on his wheelchair and kept it there for the whole session, always looking up. However, with the exception of this particular participant, we noticed that the participants were having problems with stars falling from above or at 45^{o} angle because the angles the participants needed to tilt their heads caused head and neck strains.



Arm Movement Patterns

Vertical Swipe: Arm moves up and down to catch a star.

Horizontal Swipe: Arm moves left and right to catch a star.

Extension Swipe: Angle increases at the elbow as arm moves, a downward flicking motion like catching a butterfly with a net.

Flexion Swipe: Angle decreases at the elbow as arm moves, an upward flicking motion like catching fish with a net.

Serendipitous Catch: User doesn't move to catch a star, any stars caught are coincidencidental of star spawning location relative to the arm location.

Figure 2.6: Identified Basic Arm Movement Patterns during User Studies

In summary, through video content analysis of the interviews and observations, we could argue that our game design choices were understandable and enjoyable for people with developmental disabilities, although some design elements were not noticeable by the participants, indicating the needs to revise those elements to make them more salient.

After the pilot study, we automated the video coding process by leveraging the real-time data generated during PSC game play. Through coding the data automatically, we can analyze the data in a more efficient and quantifiable way.

2.9 Revised System

The pilot study helped us tremendously in understanding how to best address many challenges in using games for rehabilitation. Another round of brainstorming sessions were performed at SDLC and Hope Services to agree on modifications to PSC with consideration of the pilot study. Additionally, we decided to automate behavioral capture to save time video coding, so that future testing could be scalable. A second iteration of the system design for PSC was done based upon the lessons learned from the pilot study.

The modifications include:

1. Speed weighting of stars were increased by 25%. The speed weighting became 50%, 100%, and 150% of user's comfortable speed for bronze, silver, and gold difficulties, respectively, to increase the varying reaction times for the user to catch stars as some users caught most of the stars. These changes were made to increase the exaggeration of difficulty between stars.

2. Haptic feedback was adjusted to emit stronger and longer pulses and a dull sound effect with the use of the strong arm. A dull bell sound is emitted whenever the user catches a star with their strong arm, in contrast to the exciting strum sound emitted when the user catches with their weaker arm. These changes were made to increase contrast between the arms.

3. The falling stars were changed to have varying sizes and shapes for the user to identify the differences between each star type. The stars were changed to bronze octahedrals, silver diamonds, and gold stars. Shade of colors was adjusted as well. These changes were made to increase contrast between star types, and were implemented due to the feedback of not noticing different star types of the users in the pilot study.

4. Data sampling rate was made customizable, which is recommended 90Hz or greater depending on computer performance, during game play and saved in a file in the following format for each datum:

    - Time [s]
    - Score
    - Last Caught [star type]
    - Weak Catches [count]
    - Weak Arm Position (x,y,z) [m]
    - Weak Arm Rotation (x,y,z) [°]
    - Strong Catches [count]
    - Strong Arm Position (x,y,z) [m]

   - Strong Arm Rotation (x,y,z) [°]
   - Headset Position (x,y,z) [m]
   - Headset Rotation (x,y,z) [°]

   The data is stored in comma-separated values format, which can be imported into MATLAB, Microsoft Excel, or other tools for analysis. Analysis can illuminate user strategy, performance, and adherence to using their weaker arm. We also created scripts to automatically calculate the trajectory and speed of each head and arm movement, the total number of stars caught with the weak arm and strong arm, and the total number of stars missed.

5. Various changes to the evaluator interface were made as well:
   - An adjustable spawn rate input was added to allow further customization for how many repetitions of movement were required of the user for the given spawn time.
   - An updated ”How to Play” menu was added with instruction text that recommended and explained input fields to teach therapists how to properly set up values in PSC and adjust the user to the virtual environment.
   - An pause menu, which is triggered by the evaluator or user, was implemented for the player view to include written instruction of PSC mechanics. This was added to ensure uniformity between researcher in how to explain the game to a participant.

The Star Catcher size and the trajectory of the stars (or rather the movement required to catch the stars) remained unchanged in the revised system. In summary, the changes described in the items above were a result of the observations made in the pilot study. With the approval and recommendation of therapists from SDLC and Hope Services, we then updated PSC and moved on to a new user evaluation with post-stroke participants.

2.10 User Evaluation

We evaluated PSC with eleven stroke survivors at Cabrillo College's Stroke and Disability Learning Center. Since 1974, the SDLC has provided an interdisciplinary educational program for adults who have survived a stroke or are living with disabling conditions. The SDLC is a unique program that starts where medical rehabilitation leaves off. Patients become students and enroll in specialized classes to develop strategies and gain skills in a supportive learning community. For the sessions at SDLC, PSC was run on a laptop in their physical exercise room. Participants were accompanied by physical therapists and caregivers.

Two participants were excluded from data analysis for reasons discussed in the next section. The remaining nine post-stroke participants consisted of five men and four women, aged 36 to 87 years old, and had been recovering from stroke for at least three years (the longest was nine years after stroke). Three participants were in wheelchairs, four used canes, and two had personal assistants to help them walk by holding their arms. Due to HIPPA privacy rule, we were not informed about their specific disabilities.

Each user was verbally instructed on the mechanics of the game, then re-read the objectives through PSC's start menu. The users were instructed to identify all logos in the environment by physically looking around in the 360° environment before initial game play, which was done to acclimatize the user to the VR experience. All users participated in an introductory round of 90 seconds of game play with Game Mode 3 (stars approaching the user from the forward view), a spawn rate of five seconds per spawn, and a base speed of three meters per second. Base speeds were then adjusted appropriately for each participant based on their physical ability and time since stroke. Next, the users were asked to try to surpass their introductory score while data was collected as an actual trial of about 5 minutes of game play. As with the pilot study, participants were interviewed after the play session. Because of the situation we observed with the pilot study, where the participants seemed to be having neck and head strains with stars that were falling from above or at a 45° angle, we only ran studies with stars approaching the users from a forward view (Game Mode 3).

2.11 Results and Discussion from the User Study

A lesson we learned from our pilot study is that, if we could record performance, strategies, and adherence in a more quantitative way, we could analyze in a more quantifiable, automated method compared to human video coding.

2.11.1 Are participants with hemiparesis able to use the controllers during game play?

Similar to our pilot study, the answer to this question was readily accessible through direct observation by researchers in the play sessions. Wrist straps and arm sleeves were useful for nine out eleven participants.

During testing, we did not have resources to reasonably accommodate two players because they were early in their recovery and unable to use their stroke-affected arm to hold the game controller even with the strap and arm sleeves. Both players tried

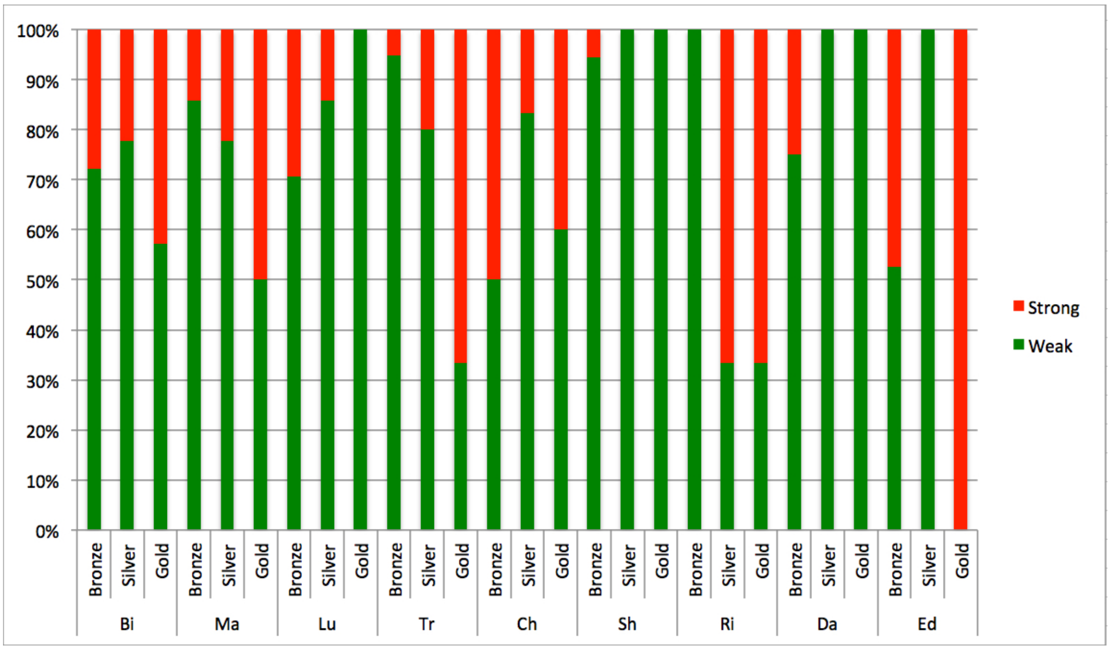


Figure 2.7: The Percentages of Stars Caught by Each Post-Stroke Participant

out Project Star Catcher using only their strong arm and both requested to participate in a follow-up study with accommodations to hold the controller with their weak arms. Data collected from these two players were skewed because they only used their strong arm to play, therefore they were excluded from analysis. Future work will include accommodation for participants who cannot grasp the controllers, our next iteration will leverage Vive Trackers to perform full body motion capture to potentially eliminate the need for a controller.

2.11.2 Can the game mechanics result in improved adherence and compliance than that of conventional mCIT?

Figure 2.7 shows the percentages of stars caught by arms with and without hemiparesis for each participant. As observed in the pilot study, with very few exceptions, participants are adherent in using their weak arms, although adherence rate went down for faster stars. Across all participants, the adherence rates for using their weak arms were 77.26%, 81.99%, and 59.31% for the bronze, silver and gold stars, respectively. This rate is higher than that of conventional mCIT [254]. The average adherence rate of 73% was also higher than our pilot study that had 65%, which may suggest that our changes to PSC had a positive effect.

2.11.3 Is the game understandable and enjoyable by our target population?

To answer whether the game was enjoyable, we relied on interview data that were coded in a similar manner to the pilot study. The participants thought the game was fun, and also noted that the experience was more fun than their regular exercise regime. Furthermore, the participants did notice that the stars were different colors, shapes, and speeds. The users did not appear to have increased difficulty with varying sizes or shapes based on observation. They also noticed the differences in vibrations and sound effects when the stars were caught with the weak versus the strong arms. Additionally, some participants clearly stated that they wished they could play longer, indicating that they really enjoyed the game.

Reactions to PSC ranged from curiosity, where they wanted to know more about future iterations of the game and new games, to enthusiasm to quickly adopt PSC for routine physical therapy. Some therapists and stroke survivors expressed interest in using PSC as a part of physical therapy routines because it is a Virtual Reality game that addresses the movements needed for rehabilitation. One therapist commented, ‘’For me, the most interesting thing and the biggest benefit for patients is the fact that the game requires the sitting balance because they have to reach and lean and come back to the center.‘’ One stroke survivor showed their excitement for VR when they explained, ‘’When you have a stroke you want to escape into another reality, and this helps you do that. I think that‘s a good thing because if you are not mobile in reality, you can escape into a world where you are mobile, and that would feel like a positive thing.‘’

It was apparent from direct observation and video observation post-session, that the participants understood the game rules and developed unique strategies to garner points and catch more stars. However, instead of analyzing the video using human coders, this time we automatically captured the strategies using our data files and scripts. It seems to us that the automatic capture was quite successful in providing us with similar data to that of human coders.

Figure 2.8 is a histogram of the head rotations of stroke survivor during game play. The X, Y, Z rotations indicated the roll, pitch and yaw movements of the heads. In this figure, we can clearly see which participants rotated their heads up and down, left and right, or remained more constant than others throughout the session. The hand and head positions are shown in Figure 2.9. We organized the participants in descending scores.

Through analyzing these figures, we could deduce that the stroke survivors were very motivated to use their weak arms, although some users (e.g. Users 3, 5, and 9) also used their strong arms extensively. In our observation, this was due to fatigue and weaknesses affecting their ability to continue using solely their weak arms, instead of attitudinal non-adherence.

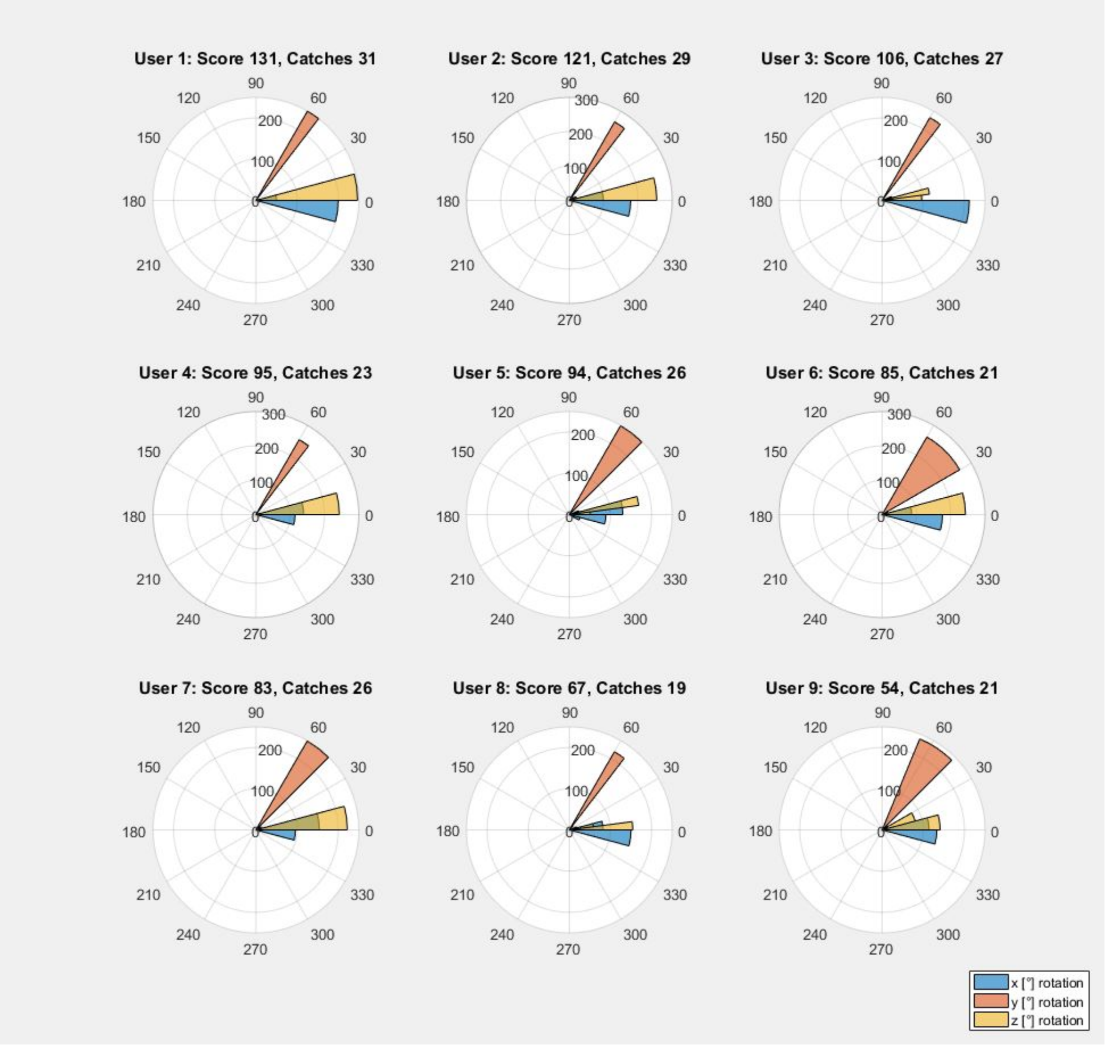
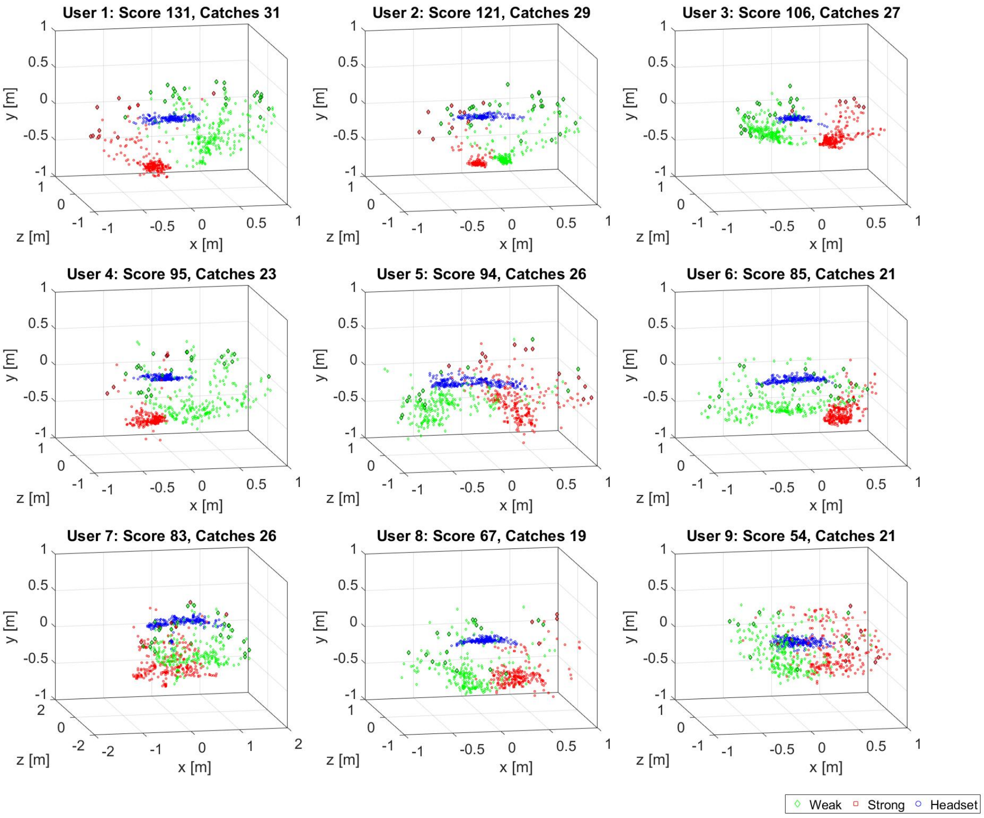


Figure 2.8: Histogram of head rotations of stroke survivors from HTC Vive



User 1: Score 131, Catches 31

User 2: Score 121, Catches 29

User 3: Score 106, Catches 27

User 4: Score 95, Catches 23

User 5: Score 94, Catches 26

User 6: Score 85, Catches 21

User 7: Score 83, Catches 26

User 8: Score 67, Catches 19

User 9: Score 54, Catches 21

Figure 2.9: Motion Capture of Hand and Head Positions. Green = weak arm, Red = strong arm, Blue = head

2.11.4 What does the data tell us about user game play strategies?

Figure 2.9 told us many things about stroke survivors' performance, strategies, and adherence. Just by comparing participants with higher and lower scores, we could see which strategies led to better scores. For example, we found the following patterns from those diagrams:

1. Those who caught more stars with their weak arms scored higher.
2. Head movements were not related to scores. Essentially, it does not make a noticeable difference in the game whether a participant moved their head a lot or not in terms of the success of catching stars.
3. Larger ranges of motions led to higher scores.
4. Strong arm movements affect the scores very little (i.e., those who moved the strong arms more did not necessarily gain higher scores).

2.11.5 Discussion

This paper reports on how we address the challenges of designing and evaluating a VR game that adapts an established physical therapy program through its game mechanics. The game was one of the first to leverage the HTC Vive for an immersive therapy experience that was evaluated by its intended users, participants with hemiparesis that need upper limb rehabilitation, as well as their therapists and caregivers. Our pilot study involved hemiparesis-affected participants with developmental disabilities. We gained useful insights on the need for simplifying the game instructions, mechanics, and reward structure, while still maintaining the idea of constraint induced therapy. Therapists from Hope Services and SDLC provided input to assure that our translation of mCIT into VR was relevant and applicable. The stroke survivors who played our revised VR experience helped verify that our updated mechanics and game design made sense.

The scoring system and VR experience of Project Star Catcher encouraged players to actively perform upper limb rehabilitative motions in an engaging and stimulating environment. User evaluations provided us with some considerations for rehabilitation games that we will share with other researchers working with similar populations or developing similar systems. The data shows that the adherence rate of the user goes down when the stars are harder to catch. This can be seen in Figures 2.4 and 2.7; there is a noticeable trend that the users ended up using their strong arm to catch the gold stars whereas the strong arm catching rate is less for silver and bronze stars. Improvement for future iterations of PSC can be to spawn more high reward stars closer to the green Star Catcher to encourage players to use their weak arm. The immersive experience made the game exciting to play, but adding different types of games will make it more interesting according to participants. Because participants developed strategies to randomly catch stars by randomly waving their arms in a swiping motion, other game types can focus on emphasizing precision and patterns in their motion.

Gathering quantitative and qualitative data enabled us to relate the performance data with the qualitative explanation of why the performance data shows a certain pattern. One lesson learned is that, at least with people with disabilities that worked with us, observation data was more meaningful than interview data. Users don't always know what they need or can express it completely, so observing them with the game gave more meaningful conclusions. This could be because the disability renders their verbal responses to be less elaborate or that they were too polite to critique our design. However, our observation data did reveal that participants enjoyed the game, either because they stated it or through our video analysis of their behaviors. Some of them also stated that they would like to play more when their session were over, which indicated that the game was not boring. We also found out that data capture and automated analysis allowed us to cut down on the time required by human coders to extract themes related to behaviors and strategies. This is helpful for scaling sample size in future user evaluations.

On the design side, a lesson learned was to keep the game simple but to design different stimuli to be more salient in their differences. For example, in our pilot study we only used colors to differentiate star speeds. This color difference was not noticed at all. However, in the second iteration when we used more than one design feature to differentiate speeds, including colors, sizes, and shapes. These were noticed by the participants in the debrief interviews.

On the strategy for effective constraint induced therapy, we learned that having larger ranges of motion for the weaker arms and varying movements (strategies) result in better game performance, which may lead to better, more effective exercise. In other words, unsurprisingly, to be more effective in doing constraint induced therapy, users need to practice a variety of movements and extend their arms as much as they can.

2.12 Conclusion and Future Work

Physical therapy can be boring, uncomfortable, and sometimes painful. This experience may result in less consistent participation from patients. We aim to contribute an engaging way of performing physical therapy through a VR game that rewards efforts in exercising hemiparesis-affected arms using the HTC Vive, a relatively new system as of this paper. Our game was developed with user-centered design method involving representative stakeholders from Hope Services and Cabrillo College‘s Stroke and Disability Learning Center.

Project Star Catcher is based on an established therapy method called modified Constraint Induced Therapy. Our study indicated that it is possible to convert physical constraint proposed by mCIT into psychological constraint through game mechanics that are understandable by our target population. Test data suggests that such users were adherent to game rules, although the adherence rate declines when the difficulty level significantly exceeds their comfort level. Our data suggests that iVR in PSC yields noticeably larger adherence rates compared to traditional forms of mCIT.

Our study provided us with design guidelines for other researchers and practitioners in similar areas:

1. Make sure that the feedback is perceived to be relevant to the goal of the therapy. Some of our users did not even notice audio and haptic feedback to indicate which hand caught the star as they were perceived to not directly relate to their goal of catching the next falling star, while our intention was that the audio and haptic feedback should inform them whether they had used the strong or weak arm so that in the future they were more compliant in using the weak arm.

2. Make the different stimuli more salient in their differences using more than one design feature. When we differentiated stars with varying difficulties by only using colors, it wasn't noticed. However, when the stars were differentiated by color, shape and size, they were noticed.

3. Always triangulate quantitative and qualitative data, especially when working

with people with disabilities that tend to be polite and might not elaborate as much as we expect.

4. Not all qualitative or quantitative data are equal, so decide which one to follow when there is contradiction - in our study we found that observation was more insightful and truthful than interview data as the participants tend to be too polite to critique our game or not as eloquent in describing their experience. We chose to weight the observation data much higher than that of the interview data in those cases.

Specifically on physical rehabilitation, our study also provided us with some design guidelines for other researchers and practitioners in these areas:

1. To maintain adherence, design the game so that it can be set, for each user, to be relatively close to his/her comfort level in terms of speed, difficulty, range of motion, etc. We noticed that even participants with the best intention of using only the weak arms switched to the strong arms when the game speed and range of motion required far exceeded their comfort level.

2. Design the game flexibly so that it can accommodate a wide range of user strategies, as long as these strategies are in line with therapy goals (e.g., if the therapy goal is to make sure that user moves his/her arms with the widest possible ranges of motions, then make sure that the game mechanics can accommodate those who prefer to swipe their arms in circular, horizontal or vertical fashions.

3. Unless it violates therapy goals, always provide game pause feature to allow for: users to rest when they feel fatigue, pick up the controller that fell, or explain the game mechanics again to the users

While this game is a step towards improving mCIT into a more engaging experience, there are future plans for the game and study designs. After exploring the increased adherence and feasibility of using Project Star Catcher in this study, we aim to perform a longer term study with a larger user group. Our experimental design was a snapshot design in a controlled environment. We plan to perform a long-term study in a naturalistic setting (in our case, either at Hope Services exercise room, at user homes or at Cabrillo College). We also aim to explore Project Star Catcher's data output into potentially integrating the use of machine learning with the physical rehabilitation process. The HTC Vive Trackers can be integrated to perform full body motion capture which will integrate the whole human body as a controller, eliminating the need to develop methods to hold the controller. This change will also enable PSC to provide full body data for therapists. Lastly, Project Star Catcher can benefit from the creation of a cloud based therapist interface to make the collection and visualization of data more impactful.

In closing, we believe that we are in the right direction in developing a VR game based on an established therapy program that is portable, engaging, and understandable while being able to provide compliance, adherence, and performance data in an efficient way.

2.13 Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 1521532 and Disability Communication Fund titled "Interactive platform for articulation and pragmatic language improvement in persons with autism/DD." 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 or Disability Communication Fund. We thank Michael Liber and all of Hope Services clients as well as stroke survivors, caregivers, and therapists from Cabrillo College‘s Stroke and Disability Learning Center for contributing to our game design and user evaluations.

Chapter 3

Project Butterfly: Synergizing Immersive Virtual Reality with Actuated Soft Exosuit for Upper-Extremity Rehabilitation

3.1 Summary

Immersive Virtual Reality paired with soft robotics may be synergized to create personalized assistive therapy experiences. Virtual worlds hold power to stimulate the user with newly instigated low-cost, high-performance commercial Virtual Reality (VR) devices to enable engaging and accurate physical therapy. Soft robotic wearables are a versatile tool in such stimulation. This preliminary study investigates a novel rehabilitative VR experience, Project Butterfly (PBF), that synergizes VR Mirror Visual Feedback Therapy with soft robotic exoskeletal support. Nine users of ranging ability explore an immersive gamified physio-therapy experience by following and protecting

a virtual butterfly, completed with an actuated robotic wearable that motivates and assists the user to perform rehabilitative physical movement. Specifically, the goals of this study are to evaluate the feasibility, ease-of-use, and comfort of the proposed system. The study concludes with a set of design considerations for future immersive physio-rehab robotic-assisted games.

3.2 Introduction

According to the latest US Census in 2010, there are more than 40 million older adults (defined as people aged 65 years old or older) living in the US, comprising 13 percent of the US population. This demographic represents a 15 percent growth compared to the 2000 US Census data [260] and is projected to continue to grow. Unfortunately, studies have shown that motor functions decline with aging [261]. The significant older population experiences an increasingly prevalent issue of motor degeneration. Age-related motor performance deficits include coordination difficulty, decreased variability of motor ability, slowing of movement, and problems with balance and gait [262]. Movement slows with aging by as much as 15 to 30 percent. Research by Seidler-Dobrin et al. suggests that older adults emphasize movement accuracy at the cost of movement speed [263]. As a result, older adults show specific deficits in the coordination of bimanual and multi-joint movements. For example, movements become slower and less smooth when older adults use their shoulder and elbow joints simultaneously as opposed to performing single-joint actions [264]. Often postural stability is also compromised with advancing age [265].

In addition to the decline in motor function, aging correlates to the progressive loss of skeletal muscle mass and strength. Frequent exercise represents an effective therapeutic strategy to augment skeletal muscle mass and improve functional performance and quality of life in older adults [266]. Many technological solutions have been researched and developed over the past decade to reduce motor loss, but there is still much to be done.

3.3 Related Work

One modern approach to address muscular impairment is virtual-reality (VR) therapy. Through the use of VR, stimulating immersive environments can be programmed to increase therapy compliance, accessibility, and data throughput [12, 13]. Psychological and physiological research has featured increasing use of VR in the prior two decades thanks to the ability to simulate realistic and complex situations critical to laboratory-based human behavior investigations [69]. Traditional forms of therapy and rehabilitation usually derive from therapist observation and judgment. Drawbacks of this traditional method are that they are often inaccurate, expensive, and timely [42]. Virtual reality, however, addresses these concerns as a useful tool for improving outcomes compared to conventional therapy by enabling accurate motion capture, telepresence based sessions, and low cost motivating experiences [72, 12]. The immersive visual capabilities of modern VR headsets, such as the HTC Vive and Oculus Rift, have had astounding promise and success with treatments ranging from exposure in Post Traumatic Stress Disorder [63, 64], Borderline Personality Disorder [65], various phobias [66, 67, 69], schizophrenia [68], and many other psychological therapies. Researchers are even reporting that integration VR into the clinical setting can reduce pain similar to the effect of analgesic treatments [61, 62].

Success in VR therapy often relates to the relationship between presence and emotion with technology's ability to bridge them [64]. Increasing the quantity and quality of stimuli in immersive VR is key to influencing user behavior and experience [70, 64]. The past five years have made strides in VR technology – VR is ever more immersive, affordable, and accessible to the average consumer with over 200 million immersive VR Head Mounted Displays projected to be sold by the year 2020 [267, 243]. For these reasons, headset-based VR systems like the HTC Vive and Oculus Rift could appeal to low-income communities. VR as a therapeutic tool has, therefore, become the most effective and affordable it has ever been and is projected by many researchers to continue along this forecasted trajectory.

Many of these studies incorporate Mirror Visual Feedback Therapy (MVFT), the visual or physical stimulation of a "pseudo" movement on the damaged limb to promote recovery [268]. Patients are given sensorimotor feedback by reflecting an abled arm in the position of the impaired arm during exercise [269]. MVFT is suggested to be a beneficial treatment for motor rehabilitation [270, 269], where clinical studies have indicated that MVFT "can serve as a versatile tool to promote motor recovery" in mobility and arm use [268].

MVFT requires the superposition of a simulated arm on a phantom limb which enables patients to relieve painful sensation and increase movement [270]. A variety of other conditions were explored with MVFT, including with stroke survivors, where a simulated limb is placed in the patient‘s midsagittal plane, thus reflecting movements of the nonparetic side as if it were the affected side to stimulate brain plasticity [271]. However, users with severe motor loss require increased physical assistance to perform MVFT, and can often require therapist intervention, or in the case of this study: a robotic wearable.

A large amount of physical therapy research has transpired on the integration of rehabilitative robotic wearable devices. Several upper-body robotic exoskeletons have been developed and explored over the last ten years with many incorporating VR. Some of these examples include the PERCRO (Perceptual Robotics laboratory) L-EXOS system [272], Rutgers CyberGlove and Master II-ND (RMII) force feedback glove [273], and Therapy Wilmington Robotic Exoskeleton (T-WREX) [274]. Through combining VR with exoskeletons that provide arm gravity support, clinical testing showed a range in improvement of mobility, strength, and satisfaction [74].

One attribute common to these exoskeletons is the use of rigid structures. A significant flaw experienced by traditional rigid exoskeletons is their inflexibility, and the burden users bear when wearing them. Devices which have few degrees-of-freedom (DoF) or heavy components inhibit some movements. This physical constraint can lead to imbalanced muscular growth and control, which can injure users of these wearable robots [275]. As a result, softer devices such as Lessard et al. exosuits have emerged as a flexible alternative to traditional rigid exoskeletons [276, 1]. This study aims to leverage such soft exosuits during VR therapy through the use of Compliant Robotic Upper-Extremity Exosuit (CRUX), as shown in Figure 3.1, to explore feasibility, ease-of-use, and comfort.

3.4 System Design

The VR experience developed, Project Butterfly (PBF), is a game that motivates users to perform upper body motion primitives by having them protect and control a virtual butterfly in a meadow while the system collects real-time data using the HTC Vive. Figure 3.3 displays an example of PBF gameplay. CRUX was integrated accordingly to provide additional tactile feedback and physical assistance.

3.4.1 The Soft Robotic Wearable

The purpose of using the CRUX exosuit was twofold: to physically stimulate movement to achieve an ideal position, and to immerse the user deeply into the VR environment. Since immersion is a crucial factor in the influence of user behavior and compliance [70], PBF paired with CRUX may significantly improve the user's rehabilitation experience.

The designed exosuit, shown in Figure 3.1, is capable of lifting the user's arm in different directions to create smooth multi-jointed movements. The concept of tensegrity for soft robotics inspired the mechanics of CRUX. Tensegrity (a portmanteau of "tensile" and "integrity") defines structures as internally prestressed, free-standing, pin-jointed networks in which the cables or tendons are held in tension against a system of bars or struts [277].

A base layer of neoprene held CRUX together. Cables are routed along the neoprene to integrate a network of "anchor points," which serve as the rigid components in an otherwise flexible system. The exact placement of the cables was determined by recording arm movement on people as they stretched out and expressed their full range of motion [1]. Through examining extensive motion capture of the arm, the area on the skin which sheared the least was determined to find the most stable places to plant anchor points [1].

Bicycle housing routes the cables onto anchor points to actuate different parts of the arm just as how tendons pull limbs. Six micro DC motors were mounted on a modular backplate and connect to the cables of the exosuit via 3d printed spool, and are manipulated through a microcontroller powering the system with a 3-Cell Lithium polymer battery. Each motor is capable of exerting 88 N of force (125 oz-in). Figure 3.1 depicts CRUX being operated by a demonstrator. The selected material of CRUX affords a compliant and lightweight design. Like similar soft exosuits, CRUX is lightweight. However, CRUX weighs 1.5 kg [276] compared to the 6.8 kg lower-limb gait-assisting exosuit developed by Wehner et al. [278] or the 2.27 kg suit by Alvara et al. [279] for upper arm force amplification

The suit's controller was designed to have the weak arm follow the movement of the healthy limb. This mirroring of limbs instigates mimetic controller design. Mirroring the movement from one side to the other side of the body was inferred from MVFT to increase motor recovery and stimulate brain plasticity [271]. Figure 3.2 depicts the CRUX being fitted to a user by an evaluator for upper arm force amplification.

To enable the mimetic control of the healthy arm onto the weak limb, wireless connectivity capability was added to connect with the Inertial Measurement Unit (IMU) networks. An IMU is an electronic device that measures and reports a body's specific force and angular rate using a combination of accelerometers, magnetometers, and gyroscopes [280]. The IMU network added to the exosuit consists of 4 IMU nodes where each node can measure 3-axis orientation of itself and then send this data back to the microcontroller. The IMU nodes on CRUX are enclosed in a 3D printed case with adjustable Velcro straps to accommodate various body sizes.

A plunger button must be engaged by the user to allow for exosuit movement. This safeguard prevents accidental actuation when the exosuit is in master-slave mode. If the user feels that the motor is performing movements that are undesired, the user can release the plunger button, effectively disengaging the motor.

In complementary locations, nodes are positioned on both arms at the lateral forearm (midway between the wrist and the elbow) and the lower medial triceps (slightly above the elbow) [1] as seen in Figure 3.2. Each node transmits pose data to the central controller to support the closed-loop function enabling the pose following from the healthy arm onto the impaired arm. It should be noted that while the VR device can perform motion capture in place of the IMUs, creating this dependency would limit the flexibility of CRUX for future use. For future example, the suit may be used beyond VR MVFT to assist with active daily living activities, where the controls and level of assistance are calibrated during the VR therapy sessions.

3.4.2 The Immersive Virtual Reality Experience

A motion primitive is defined as a distinct movement achievable by a single joint which creates a unique degree of freedom (DoF). Thus, upper body motion primitives can be thought of as indivisible building blocks that can be combined and permuted into a broader range of potential movements. The HTC Vive, one of the highest grossing VR Entertainment Systems [267] developed by the Valve and HTC Corporation, can be used as a powerful tool to both track these motion paths and motivate the user to achieve these motions. The Vive is a VR Head-Mounted Display that implements room scale 4x4 m outside-in tracking technology by utilizing a “lighthouse” system of lasers which enable the user to interact with the virtual environment through accurate motion capture in a 3D virtual space [257]. Complementary to the HTC Vive are two handheld controllers that feature dynamic haptic feedback which enhances spatial orientation [257]. The HMD provides a 110-degree field of view and 90 Hz refresh rate [258]. Motion capture is tracked at 120 Hz using infrared laser sweeping and photo-diodes that enable for recovery of position and orientation [257]. The HTC Vive also features a safety guidance system preventing users from potential injury in the real world environment [258, 259]. The worst case tracking jitter of the system has been reported to be under 2.1mm with an accuracy of an absolute 2mm error [255]. Resultingly, the HTC Vive allows for the ability to extract accurate gameplay data while providing an enveloping experience of touch, sound, and sight.

Paired with an HTC Vive controller, the exosuit assists the user during VR gameplay. Testing of the system targeted two pairs of motion primitives: elbow extension/flexion, and shoulder abduction/adduction. Biceps received assistance by replacing the user's CRUX supported HTC Vive controller as a bubble shield and having them protect the butterfly from incoming rain and projectiles through a therapist-specified customized range of motion path. Haptic feedback is enabled so that the user is indicated with strong pulses whenever the motion primitive was not followed (failure to encapsulate the butterfly inside the bubble). To increase the incentive and track compliance, the user receives a scoring point per every half second that they mirrored the required motion primitive. This multi-sensory feedback guided users according to the objective of the game.

To generate the environment and the mechanics of PBF, the Unity v2017.1.0b4 Game Engine along with the SteamVR Unity plugin v1.2.0 became the chosen development tools. Both Unity and SteamVR hold a large amount of open-access documentation, including flexibility with programming languages such as Javascript and C# [281]. Using Unity's built-in physics engine, the Rigidbody class was used to model the butterfly along with spherical colliders that detected contact with the butterfly. Assignment of the moving projectiles and the butterfly with the rain were set to a time-dependent spatial state, allowing for global physics-based events to influence data capture. Runtime data collection was captured using C# and Microsoft .NET Framework at speeds of 90Hz and higher.

To assist the user evaluation process, PBF includes a dynamic evaluator GUI, seen in Figure 3.4, which automatically prompts the evaluator or therapist to measure the length of each participant's arms through the motion capture (measuring the x-z plane maximal distance between the VR HMD and the Vive Controller) or manually entered ranges. This calibration stage is achieved entirely in Unity during run-time with the HTC Vive so that user's who may not have access to CRUX in the future can still play PBF without physical assistance. Figure 3.4 also features the option to change the repetitions per minute of each motion primitive, and data exportation rate of each game session. In short, the evaluator GUI allowed the evaluator to tailor the game to each unique individual and customize data throughput.

The game-themed goals became protecting a butterfly from heavy rain and projectiles using a bubble (as the avatar for the controller of the weak arm), which focused on bicep curl and lateral arm raise exercises. Users are instructed place to the bubble around the butterfly to achieve a high score, and to "protect the butterfly". These mechanics required the users to smoothly follow the flight of the butterfly within +/- 0.1m of the required motion path. As a result, the biceps and lateral arm raise minigames became scripted movements of high accuracy and scripted timing repetitions when a user performed a motion primitive.

Lastly, a tutorial started before every game that the evaluator could enable to adjust the user to the VR environment. In the tutorial, users were asked to identify three images placed on their left, on their right, and behind them. Identification in the 360 VR environment was implemented to familiarize the user with virtual reality and show them that they can look in any direction. Additionally, a “How to Play” menu was added to teach evaluators how to test users, and in-game instructions were added to clarify the goal of each minigame. This pre-gameplay stage simultaneously introduce users not only to virtual reality mechanics (such as 360-degree views) but also to PBF specific mechanics like the arm movements required to earn a higher score. These designs were intended to tailor the game to each unique individual smoothly and intuitively, all while increasing the incentives to perform the objective of the game.

3.5 Usability Study

The study investigated the usability of PBF synergized with CRUX system design, which includes ease of use, comfort, and set baseline feasibility with nine elderly users having motor dysfunction. Users were observed playing the virtual reality game with the exosuit as they performed tasks that identified the ability of targeted muscles and muscle groups. The nine elderly participants in usability evaluations represented three patient segments:

1. Three retirees from Elderday Adult Day Health Care Center, Santa Cruz, CA, represented a mental disability use case because they were affected by memory loss or dementia. Evaluations with these users were conducted at Elderday.
2. Three stroke survivors recruited from Cabrillo College’s Stroke and Disability Learning Center (SDLC), Capitola, CA, represented a physical disability use case because they were affected by neglect syndrome, meaning the motor functions of one side of their body was impaired by their stroke. Evaluations with these users were conducted at SDLC.
3. Three retirees living independently within their community, outside of daycare or

hospice, which do not have a significant physical or mental disability. Evaluations with these users were conducted at the University of California, Santa Cruz.

All participants ranged between the ages of 60 to 80 years old and were previously unfamiliar with both virtual reality and wearable robotics. These user groups represent three demographics: physical disability, mental disability, and no disability-were selected because the physical and mental disabilities are likely the target demographics for PBF and CRUX. This preliminary study serves as feasibility to justify further studies with larger user group sizes and gain design insights.

For each user, the evaluation began by an proctor giving a detailed explanation of what PBF and CRUX are, answering questions as needed. Then the user was given a tutorial period where they walked through the "How to Play" menu and played the game without being recorded. The tutorial consisted of playing two rounds of the biceps minigame and the lateral arm minigame for one minute each. The first round with the minigames was played without wearing the CRUX exosuit and the second was played while wearing it to allow for MVFT. This allowed the user to try out the controller and learn PBF's basic game mechanics. The point score from the tutorial period helped the evaluator calibrate each user (adjusting arm length and speed).

After the tutorial period, the test commenced with the user playing recorded sessions while wearing the CRUX exosuit and playing 1-minute minigames, shown in Figure 3.6. Sessions were recorded using a webcam, collecting gameplay scores and positional data, and conducting post-session interviews which allowed for the collection of quantitative and qualitative data. User evaluation interviews were executed by asking users questions pertained to their experience from a prepared form with questions such as:

- Open-ended questions:
  - What day-to-day tasks do you struggle with?
  - How many years have you been doing physical rehabilitation therapy?
  - Do you enjoy video games?
  - How can the virtual reality game be improved?
  - How can the exosuit be improved?
- Rate the statement on a 5-point Likert scale:
  - **Q1:** My current therapy is engaging.
  - **Q2:** The virtual reality game was enjoyable.
  - **Q3:** I became fatigued while playing the virtual reality game.
  - **Q4:** The virtual reality game distracted me from pain when doing physical movement.
  - **Q5:** If I had access to virtual reality therapy games, I would use it in the future.

After evaluations, proctors reviewed gameplay and interview data and commented additional thoughts or observations. Note that resulting answers from **Q1-Q5** can be seen in Figure 3.5.

3.6 Results and Discussion

The three primary objectives of the study were to evaluate the baseline feasibility of the system, ease-of-use of the system, and comfort of the system. These objectives are judged according to quantitative metrics obtained from recorded metadata of users while playing PBF. Qualitative data was acquired during post-gameplay interviews using mixed 5-point Likert Scale questions and open-ended questions. The Likert scale questions are summarized in Figure 3.5.

3.6.1 System Feasibility

Most participants felt that CRUX augmented their upper limb movements to some degree. Specifically, one participant (see User 1 in Figure 3.7) mentioned that their arm movements became more “effortless” with the use of CRUX. Specifically in raising their arm laterally, which had been challenging for them before donning CRUX. They further elaborated that it was the first time in weeks that they did not experience painful throbs when raising an arm above their shoulders.

To observe an example of such movement, the position of the CRUX assisted controller and controller target, the butterfly avatar, were graphed against four scored users from SDLC (Users 2 and 4) and Retirees (Users 1 and 3) for the lateral shoulder raise minigame as seen in Figure 3.7. All four CRUX supported users were able to achieve a score of 111/130 or higher, indicating that these users were compliant to the motion primitive path for over 85% of gameplay. Additionally, the speed of the user’s controllers, headset, and butterfly avatar (target to follow) is graphed as shown in Figure 3.8 as well as the acceleration of the user’s weak arm controller shown in Figure 3.9. Each of these users demonstrated significantly different gameplay movements even though they had close scores. When considering Figure 3.7, User 1 had sharp changes, User 2 had shaky movements, User 3 had smooth moves, and User 4 had smooth and shaky movements. When looking at movement speed in Figure 3.8, Users 1 and 3 maintained a constant speed, indicating a greater control of their weaker arm. Whereas Users 2 and 4 had significant spikes in movement speed indicating a lack of control in their arms. It should be noted that while the exosuit assists in achieving position, it does not reduce the shaking of the limbs. Figure 3.9 reflects the findings in Figure 3.8 through acceleration, where User 1 maintains the most control nearing almost no acceleration, followed by User 2 who spikes but nears zero, and Users 3 and 4 experienced occasional large shaking. Despite the difference in fine-motor strength and precision amongst participants, each user was able to achieve the desired goal of protecting the virtual butterfly. This might suggest that PBF can be accommodating for people at various stages of their physical therapy, thus making it more accessible. Also, while User 1 and User 4 have differently lengthened arms by about 0.2 meters, the compliance of over 85% can be seen visible as their green arm paths overlay the object red path arch in Figure 3.7. Although these results are promising, the sample size is not statistically significant, which warrants further study as noted in the Limitations Section 6.5.

On the subject of assistance, most participants felt that CRUX affected augmenting their limbs. However, they all asked for stronger motors on the exosuit. When asked for a potential reason to this, most users indicated that they felt that the motors were not powerful enough to make the difference that they were expecting. One user (see User 3 in Figure 3.7) commented that they understood the function of the device but “[wanted] even more power behind it.” Additionally, none of the participants felt that the exosuit made it more difficult for them to move.

3.6.2 Ease of Use

An interesting observation was that users said they "knew the goal was always to protect the butterfly," which was a gameplay theme added when brainstorming ideas with potential users. Creating in-game goals which center around an archetypal emotional response may have generated a quicker understanding of gameplay mechanics in the evaluated users. When asked what they thought of the difficulty, almost all participants thought it was appropriately challenging. This makes sense as the game‘s difficulty was adjusted based on their first gameplay before and during the tutorial period.

By automating all control and dynamic elements of both CRUX and PBF, there was an increase in compliance when performing motion primitives as well as reducing the contention between the suit and virtual reality. Users tended to agree that they noticed the exosuit was assisting them smoothly as if part of the gameplay, due to the mimetic control, suggesting that automating CRUX user-input may have benefited players immersion.

3.6.3 Comfort

Users asked for greater ventilation in CRUX, due to the form-fitting style of the neoprene. Additionally, since CRUX was prototyped using commercially available wetsuit neoprene, not all users equally fit the base layer. The majority of the participants still responded that they enjoyed playing the minigames with CRUX as seen in Figure 3.5.

Through observation and commentary, no one seemed displeased with the aesthetics of the minigames. Some users complimented the aesthetics, citing the goal of butterfly protection as “fun,” “engaging,” and even “meditative.” This is particularly exciting news since many people unfamiliar with video games are often intimidated or dissuaded from immersive virtual environments, especially considering that five out of nine participants responded that they dislike video games. The user who mentioned that the games were “engaging” initially discussed their reservations towards video games and how they preferred “real things.” When users are psychologically comfortable with a system, they are more likely to benefit from it.

From these observations and recordings, we have found that most users believe PBF was useful in helping them move their arms as seen in Figure 3.5. Combining CRUX’s master-slave system and PBF’s butterfly protection mechanics actively encouraged players to perform the required movements while being distracted from their physical therapy as agreed by six out of nine users (and one additional neutral response). These users found PBF enjoyable and agreed that the game did distract themselves from the exertion of physical moving. Seven out of nine users stated that they would regularly use PBF & CRUX in the future for exercises if it were available to them.

3.6.4 Discussion

Both the quantitative and qualitative data generated suggested that in the short term, the system can help users achieve in the moment tasks by augmenting users’ upper body strength to enable them to move their upper limb more easily. In the long term, the paired-system serves as a boost to help them train their weaker upper arm and make it stronger over time. In the case of stroke survivors from SDLC who suffer from neglect syndrome, it additionally helped them in recognizing what independently driven movement on the neglected side feels like again. Qualitative lessons learning suggest the following:

- The exosuit must afford independence on behalf of the user. In the gameplay sessions, the experimenters helped users to don and off the suit. If this suit is indeed to be used as a home-based exercise system, users must be able to put on and take off the suit by themselves, which at the moment is still challenging.
- Using a neoprene top as the base to CRUX limited those who could wear it and made it hot when worn over an extended period and difficult to adhere. In future iterations, exosuits like CRUX needs to be more easily customizable, for example by replacing neoprene in the exosuit with elastane (e.g., Lycra).
- Future iterations of robotic support should make the augmentation slightly more powerful, so the aid from CRUX is more apparent to users.
- Given that the PBF and CRUX are new to players, they must be acquainted with the technologies individually. When given even just a few minutes to become acquainted, players made much better use of the CRUX when playing PBF and were less overwhelmed.

Through the study, a set of design guidelines was compiled for other practitioners of wearables and VR games to augment upper limb movements and motivate exercises, especially in older adults and people with motor impairment:

- Exosuit needs to be designed to be more flexible to fit a variety of body sizes and shapes.
- Exosuit augmentation needs to be noticeable, perhaps by adding an in-game UI element such as the bubble lighting up when the suit is activated. This will

hopefully keep the user immersed while also providing exosuit control indication to the user.

- Reduce various forms of stimulation to the minimum without affecting the exercise goals without breaking immersion. Feedback and color variations went unnoticed by some of the users as they were not related directly to their goals.
- Stay within the appropriate difficulty of the users in terms of game speed, ranges of motion, etc.
- The user feedback suggests that the users enjoyed protecting the butterfly, as it caused an "emotional attachment." Emotion-driven immersion suggests a powerful tool in creating engaging experiences.

3.6.5 Limitations

One user responded "Disagree" for questions **Q2** (Enjoyable Game), **Q4** (Distracted from Pain) and **Q5** (Future Use) from Figure 3.5. The user commented that the novelty of VR was fun, but continuously protecting a butterfly would become boring if they used it for regular therapy. The user requested that future games should have a compelling storyline to keep them interested and motivated. The intent of PBF was preliminary feasibility, and future games should be developed and studied based on lessons learned, trying out game mechanics, and the targeted therapy desired. Concerns that this user poses should be investigated in subsequent games to determine if specific game mechanics are more preferred by users than others.

Only three motion primitives were explored and converted into VR through PBF. Specifically, they were the Lateral Shoulder Raise, Forward Arm Raise, and Horizontal Shoulder Rotation. Future studies should incorporate a more variety of motion primitives to catalyze potential benefits to the users' potential mobility improvements. There may also be potential in incorporating common assessment tools into the VR environment for automated assessment such as the Apley's scratch test (Shoulder Mobility) [282], Wolf Motor Function Test (Upper-Extremity Mobility) [283], Fugl-Meyer Assessment (General Motor Ability) [284]; all of which assess motion primitives for active daily living.

Furthermore, this study would benefit from a larger sample size of the users, more therapists could be involved, and more stimuli and testing must be done to determine feasibility, compliance, and design further. There is a possible novelty effect of the VR, where most of these users were exposed to the VR for their very first time. The results suggest that the game design may account for the novelty effect. However, a long term study must be done to address these possibilities adequately. Subsequently, a long term study is being planned with local hospitals in Santa Cruz, California, to explore the effects of Project Butterfly with CRUX further. An IRB protocol for such is currently under review for approval.

3.7 Conclusions and Future Work

Project Butterfly reports on the design and evaluation of a unique VR experience paired with a soft body robotic wearable exosuit. The pair of these technologies have been developed as a novel experience to rehabilitate upper-extremities. CRUX reduces the burden and rigidity experienced by users of traditional wearable robots through its softer, more structurally compliant constitution. However, a material other than neoprene should be used to make it more comfortable and less likely to overheat the user. To tailor to a more significant number of DoF, the designed VR game, PBF, is aimed at focusing on motion primitives expressible in soft exoskeletons – actions which the healthy arm can perform that can be combined and permuted into all upper-extremity movement. PBF thus serves as a motivator for the user to complete virtual objectives and consequently, actual motions. The completion of these objectives is assisted through CRUX's augmentation of their upper body strength to perform the game-specific movements. When evaluating the baseline feasibility of PBF and CRUX in augmenting and promoting proper arm movement as defined by the established motion primitives, most users were able to complete appropriately challenging arm movements, suggesting that PBF and CRUX gave users suitable strength of their augmented arm. Additionally, the system's ease-of-use and comfort were analyzed, and most users felt that they were confident capitalizing on the therapy system.

Virtual reality paired exosuits could prove useful to make engaging therapy for users with upper-extremity impairment. For more significant impact, designing a new exosuit peripheral and increasing more types of accompanying VR minigames which augment various muscle groups (i.e., pectoral muscles and dorsal muscles) can further improve the rehabilitation. These muscle groups support upper-extremity actions and strengthening them could bolster arm muscles as a result. In a similar vein, using new materials for a future iteration of a soft exosuit focused for VR could make this technology more comfortable and accessible.

Furthermore, the real-time data produced from the HTC Vive and Vive Tracking units can be integrated with the exosuit. Therapists and users may also potentially benefit from further data extrapolation with the HTC Vive. A complete view of a user's body, achievable through more precise motion tracking and inverse kinematics, and could identify confounding postural issues, such as slouched backs and other movement biases, which a physical therapist would want to be aware of during gameplay. A long term study is being planned with local hospitals to explore the effects of Project Butterfly with a next-generation CRUX design that accommodates more body types, sizes, and weight. Finally, with the plethora of positional and behavioral data output produced from this VR experience, there is potential to integrate machine learning protocols and AI to optimize suit controls and game difficulty to improve rehabilitation results.

The baseline feasibility, ease of use, and comfort created by synergizing an immersive physio-therapy VR game with an actuated soft robotic exosuit had promising results in the potential future of a more accessible, affordable, and personalized rehabilitation. More research is needed to expand upon the preliminary work of this study, discover best practices of soft exosuit integrated VR, and validate clinical utility. Subsequently, there are more butterflies to follow on the path ahead.

3.8 Acknowledgements

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 would like to thank the following UC Santa Cruz affiliates for their support: John McIntyre, Pattawong Pansodtee, Ocean Hurd, Dustin Halsey, Evanjelin Mahmoodi, Noah Brown, and Dr. Daniel Goodman Shapiro. The authors additionally thank the participants of this study for their time and effort.

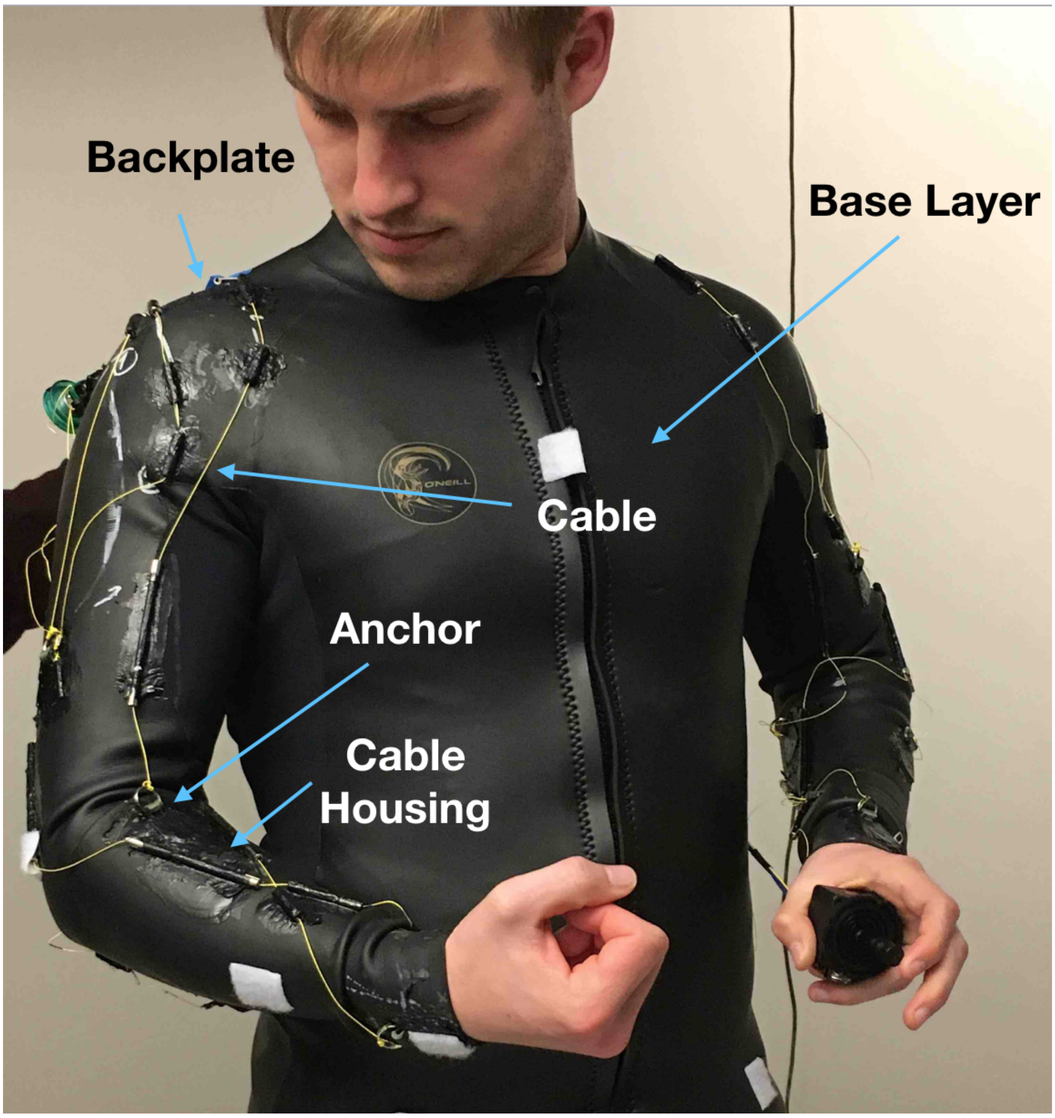


Figure 3.1: A demonstrator wears CRUX [1] (without IMUs). CRUX is an augmentative wearable soft robot for upper-extremity rehabilitation and can be combined with VR through Project Butterfly to enable immersive rehabilitation.

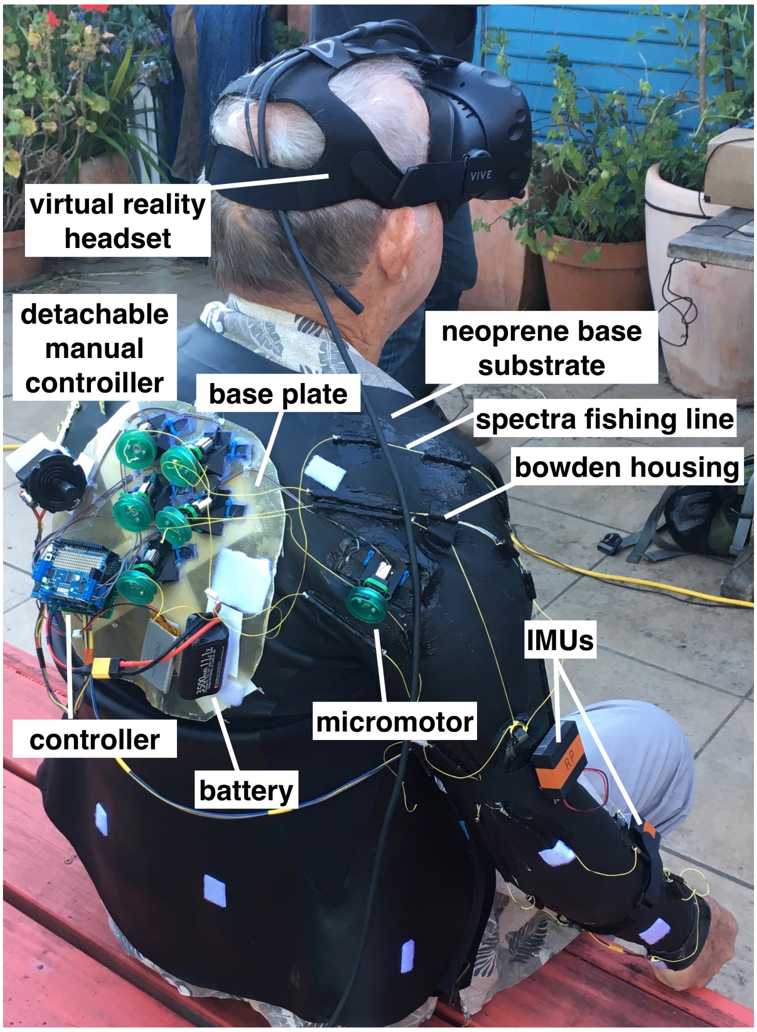


Figure 3.2: A participant exploring CRUX [1]. Control is achieved by using IMU Nodes and an internal controller for leader-follower mimicry. A user can control their impaired arm using their healthy arm to match the movement path.

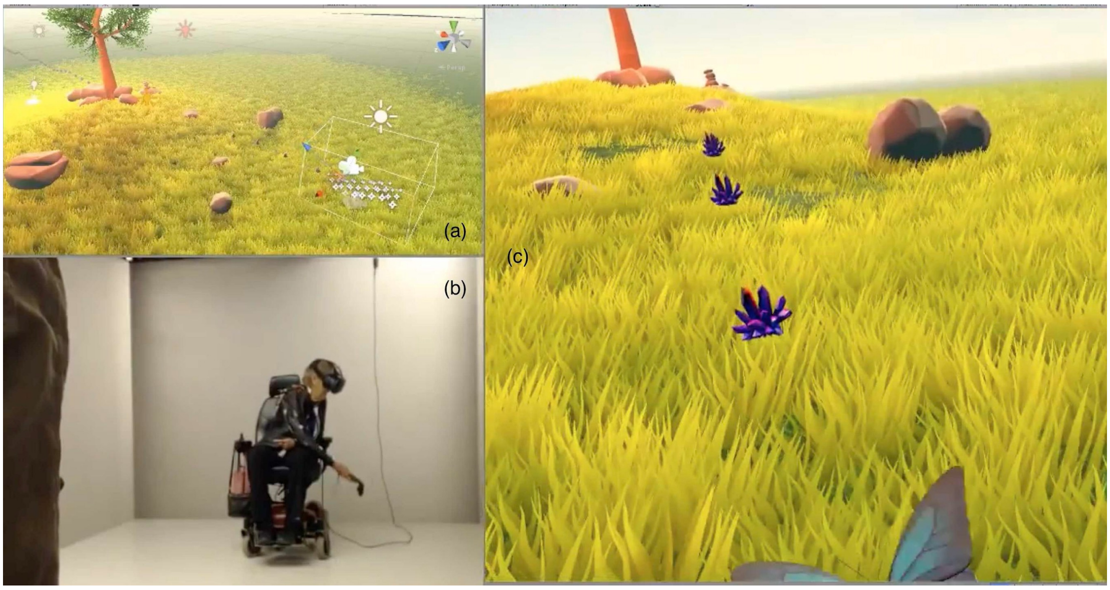


Figure 3.3: A user playing Project Butterfly. a) is the study proctor's on-screen view b) is the in-person view of the study proctor and c) is the in-game view from the user's perspective.

Figure 3.4: Dynamic Project Butterfly Evaluator Interface. This UI and the data it gathered allowed for more balanced games for each subsequent participant.

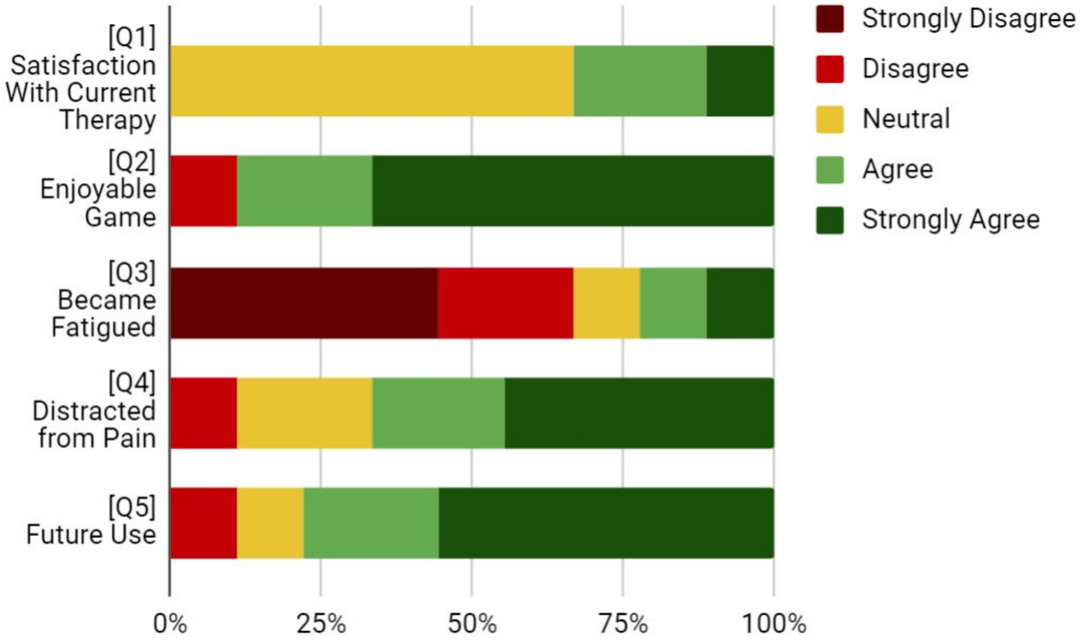


Figure 3.5: Nine user responses of 5-point Likert Scale Questions pertaining to PBF

Figure 3.6: A user performs lateral arm raises during a Project Butterfly game session with CRUX. To boost their scores, users had to mimic the flight path of the butterfly, which in this case was an up-and-down motion most easily copied by raising one's arm up-and-down similarly.

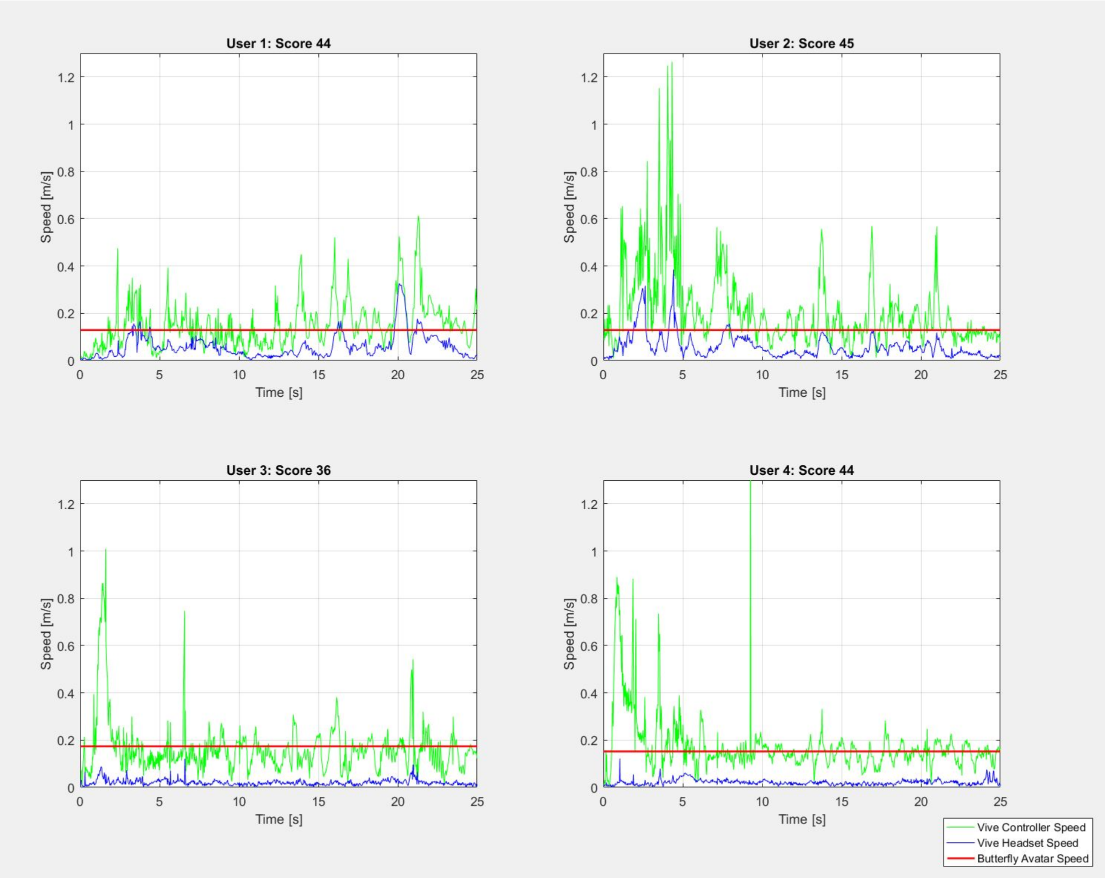


Figure 3.7: Position of the CRUX assisted controller during the lateral arm raise minigame. These graphs depict four users attempting to mimic the movement of their targets (the butterfly avatar). Red is the butterfly, green is users‘ movements. Red arcs which are closely matched by green trajectories (as shown in these graphs) mean that users are successfully completing in-game objectives, which are tailored to test their physical limits when wearing the exosuit.

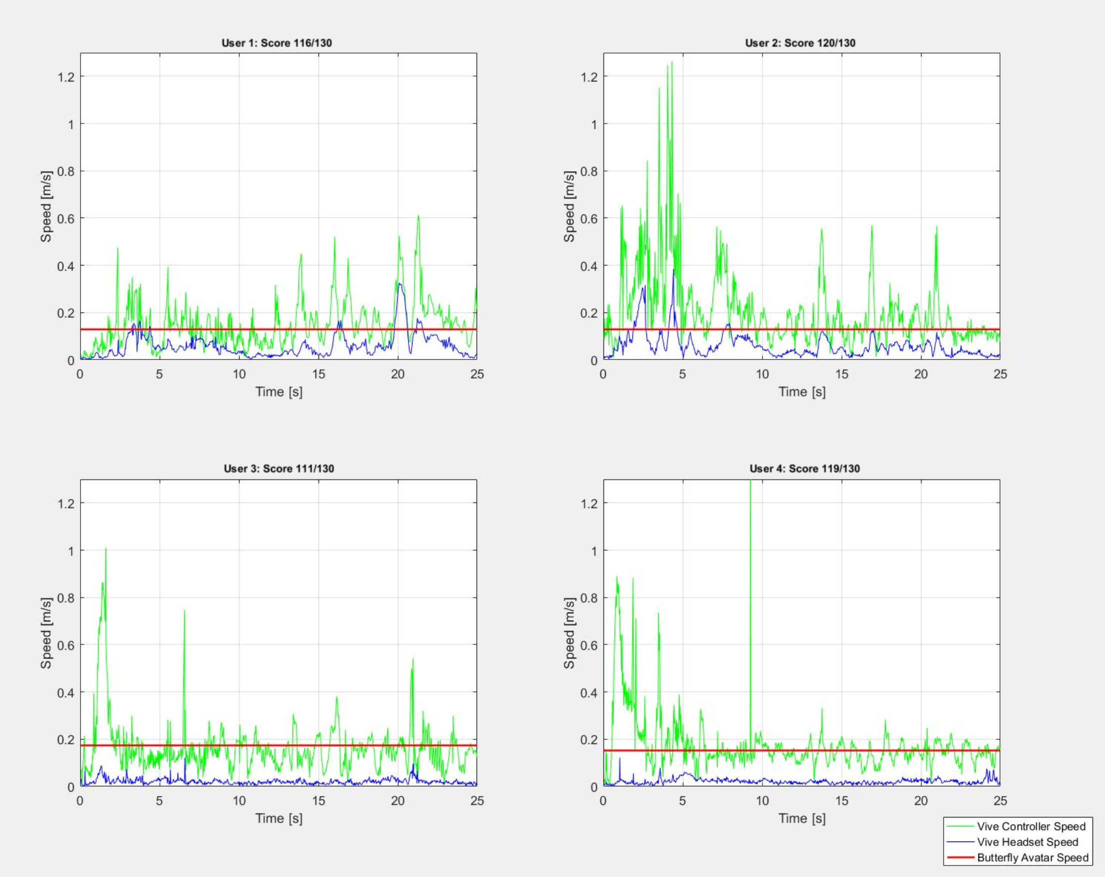


Figure 3.8: The speed of the user's handheld controller and headset as they attempt to catch the virtual butterfly avatar during the lateral arm raise mini-game. (Green is the speed of the Weak Arm Controller, Red is the speed of the Butterfly, Blue is the speed of the headset)

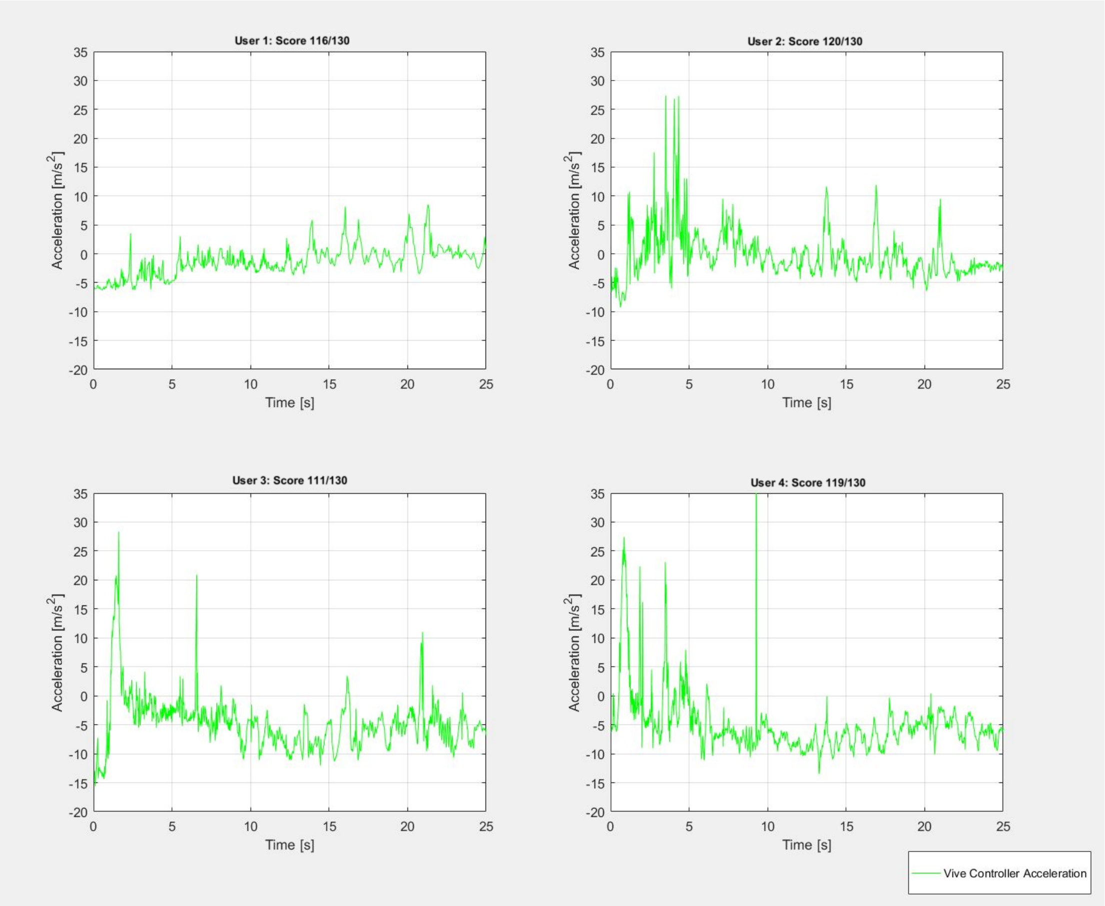


Figure 3.9: The acceleration of the user's handheld controller as they attempt to catch the virtual butterfly avatar during the lateral arm raise mini-game. Users with high scores and large fluctuations in acceleration are able to react quickly to in-game obstacles, suggesting a higher level of control and strength than those who cannot. (Green is the acceleration of the Weak Arm Controller)

Part III

EXPLORING IMMERSIVE VIRTUAL REALITY EXPERIENCES FOR EMOTIONAL INTELLIGENCE

Chapter 1

Introduction

This part investigates how immersive virtual reality experiences can be personalized for greater emotional intelligence from [18, 23]. We demonstrate how affective models (e.g. the Pleasure-Arousal-Dominance model of emotion) can be translated into immersive virtual reality experience design through artificial intelligence and user co-design through a variety of visual and haptic stimuli. These application studies end with design considerations for future researchers interested in utilizing adaptive affect models within their virtual experiences.

Chapter 2

iSAM: Personalizing an Artificial Intelligence Model for Emotion with Pleasure-Arousal-Dominance in Immersive Virtual Reality

2.1 Summary

Emotion, a crucial element of mental health, is not often explored in the field of immersive Virtual Reality (iVR). Enabling personalized affective iVR experiences may be incredibly useful for the expansion and evaluation of serious games. To further this direction of research, we present a playable iVR experience in which the user evaluates the emotion of images through an immersive Self-Assessment Manikin (iSAM). This game explores a pilot system for enabling efficient online fine-tuning of a user's Pleasure-Arousal-Dominance (PAD) emotional model using personalized deep-learning. We discuss adapting the International Affective Picture system (IAPs), in which our

Artificial Intelligence (AI) model responds with a personalized image after learning from ten user supplied answers during an iVR session. Lastly, we evaluated our iVR experience with an initial pilot study of four users. Our preliminary results suggest that iSAM can successfully learn from user affect to better predict a 'happy' personalized image than the static base model.

2.2 Introduction

At this time, iVR Head Mounted Display systems (HMDs) have garnered wide commercial adoption, with over 200 million projected headsets sold since 2016 [52]. In addition to entertainment purposes, iVR holds vast potential for serious games, which has been on the rise due to the benefits of programmable iVR for physical and cognitive applications [69]. High-fidelity motion capture and telepresence capabilities allow these assitive experiences performed in VR environments to increase user game compliance, accessibility, and data throughput while using commercially available components [72, 13, 12, 234]. In this paper, we explore how iVR can be utilized to personalize an experience for emotion using the Unity3D Game Engine and the HTC Vive iVR System. The system personalization was all performed during runtime to create a unique situation for every user. This research may be of interest to interdisciplinary researchers at the intersection of immersive media, artificial intelligence, and healthcare intervention.

In terms of modeling emotion, Paul Ekman describes nine principles for basic emotions. Ekman argues that: emotions have universal signals, are found between animals, affect physiological systems, are triggered by universal events, are coherent, have rapid onset, brief duration, are appraised automatically and subconsciously, and they are involuntary [99]. These principles provide a theoretical framework for quantifying emotions and starting empirical studies on affective states. Subsequently, many researchers have explored how to quantify these basic emotions (anger, fear, sadness, enjoyment, disgust, and surprise) in media such as music and photos [100, 101]. Considering these kinds of works, how might an emotional state be directly mapped through an iVR environment that can also provide coherent and fast responses to user interactions?

To answer this question we chose to utilize the Pleasure-Arousal-Dominance (PAD) emotional model, a conceptual construct explaining that human responses to environments can be quantified in terms of three independent bipolar dimensions [285, 286]. These dimensions of PAD can describe the emotional response from environments through pleasure-unpleasant (P), arousal-unaroused (A), and dominant-submissive (D) states[286]. To measure PAD from our users in an iVR environment, our project uses the Self-Assessment Manikin (SAM), one of the most widely used surveys for evaluation of emotional states. SAM allows for quick, non-verbal, culture-free, and language-free retrieval of PAD response to a given stimuli [287]. Applications of SAM include the University of Florida's Center for the Study of Emotion and Attention (CSEA) affective databases for pictures, audio, and words [8]. The CSEA database contains statistically-based media-to-affective-value models and has been explored with Event Related Potentials, functional Magnetic Resonance Imaging, Pupil Dilation, and more [7, 288, 107, 108]. For example, Waltemate et al's avatar personalization study utilized SAM to evaluate emotional and social experience in response to presence and immersion in embedded user avatars in an iVR environment [106].

In terms of creating adaptive experiences with PAD evaluation, i Badia et al. employed biofeedback to infer affect using the International Affective Picture system (IAPs [8]) while users navigated a virtual maze [168]. This study employed a post-test SAM, but did not query subjective feedback of self-perception from users during gameplay. With considerations of the works discussed in these sections, we sought to create an experience that could learn from the user's PAD response in a runtime iVR game. To this end, we decided to employ SAM as a runtime input mechanic to enable appraisals of user emotional states during an interaction, this then informs the prediction of our AI. Translating the PAD emotional model into a runtime mechanic for game engines may yield immense potential in an iVR environment. This work will lead to a model that can help personalize emotional engagement which could lead to more effective experiences for a variety of serious game applications in health, rehabilitation, and entertainment [69].

2.3 System Design

This project, which we dub iSAM (the immersive Self Assessment-Manikin), leverages the capabilities of immersive VR and AI to create a playable experience that learns from the emotional response of players. Specifically, there were three tasks:

- Incorporate the SAM Pleasure-Arousal-Dominance emotional model into an iVR experience [288, 7].
- Establish a methodology of dynamic learning from user emotional responses for adaptive affective models.
- Evaluate the effectiveness of said methodology in dynamically adapting affective models.

We explored several approaches in our prototype to understand how emotion may be factored into curated iVR stimuli, such as a rule-based heuristic model, a random forest model, and a ResNet-based deep learning model. The prototype uses images from the IAPs, but we structured it to allow extension into domains beyond images – for example, sound or aromatics. We log user responses during runtime for evaluation and potential reuse in other projects personalizing iVR stimuli through emotion. The potential of constructing new affective databases could be useful, especially in domains where the IAPs are lacking, such as aromatics, 3d audio, and haptic feedback.

2.3.1 Interaction Design

In our iVR experience, the user enters a room which is semi-enclosed and floats in an empty neutral space. A picture frame is in front of them. The user is told that the picture within the picture frame contains a "lost memory" of a virtual character, "Sam." By responding to each of the images presented to them using the SAM affective rating scale [7], they are helping Sam recover a lost happy memory. The user inspects an image in the picture frame for 12 seconds, after which a screen appears on their left wrist, showing a SAM affective rating scale. The user evaluates the image and submits their PAD response. After this, the picture frame updates to a new image, which the user again inspects for 12 seconds before rating. The interface used by the evaluator is illustrated on the left side of Figure 2.2. The user's view of the SAM interface and a third-person view of the iSAM environment are shown on the right side of Figure 2.2.

Each of the responses of the player allows the AI environment to update its model of image-to-affect corresponding from the baseline IAPs-based model to a more personalized responses. After appraising ten pictures from the IAPs, we show the player one last image, which we present as a lost and rediscovered happy memory of Sam. The flow of the interaction for the user is shown in Figure 2.1, and the underlying logic

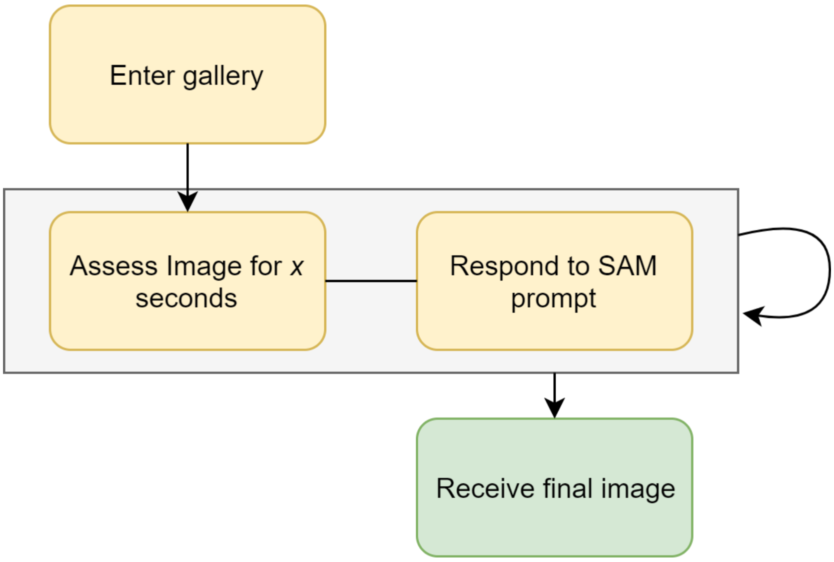
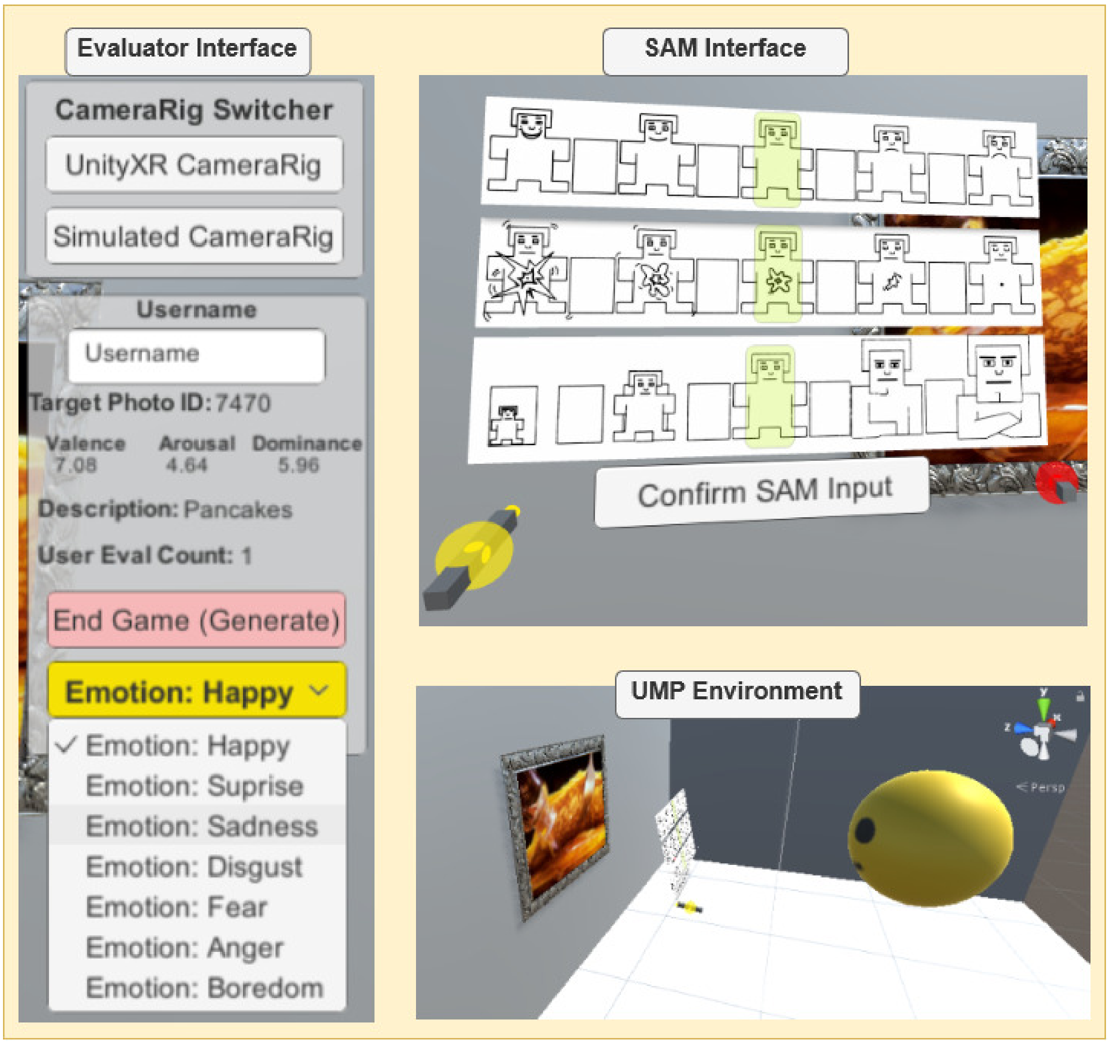


Figure 2.1: iSAM user interaction flow.

driving this experience can be seen in Figure 2.3.

2.3.2 Implementation

We designed iSAM using several components: the Unity3D Game Engine, a python server for affective models including machine learning models, and two databases: the IAPs database and our logfile system database in which user data was stored during runtime ("iSAM database" in Figure 2.3). We utilize these databases to handle the retrieval of IAPs baseline PAD data into the iSAM experience, the retrieval of runtime user responses, user response logging, and user emotional model inference. For each round of user interaction, the "Get USER PAD" component retrieves an image from the IAPs database and waits for user response. The IAPs database consists of pairs of images and metadata, which includes image category, as well as the pleasure, arousal, and dominance factors of each image with the standard deviation for each value. There are three versions of baseline metadata from the IAPs database: one with data from



Target Photo ID: 7470

Valence	Arousal	Dominance
7.08	4.64	5.96

Description: Pancakes

User Eval Count: 1

Figure 2.2: iSAM evaluator interface (left), Input Interface (top right), and Gameplay Environment (bottom right).

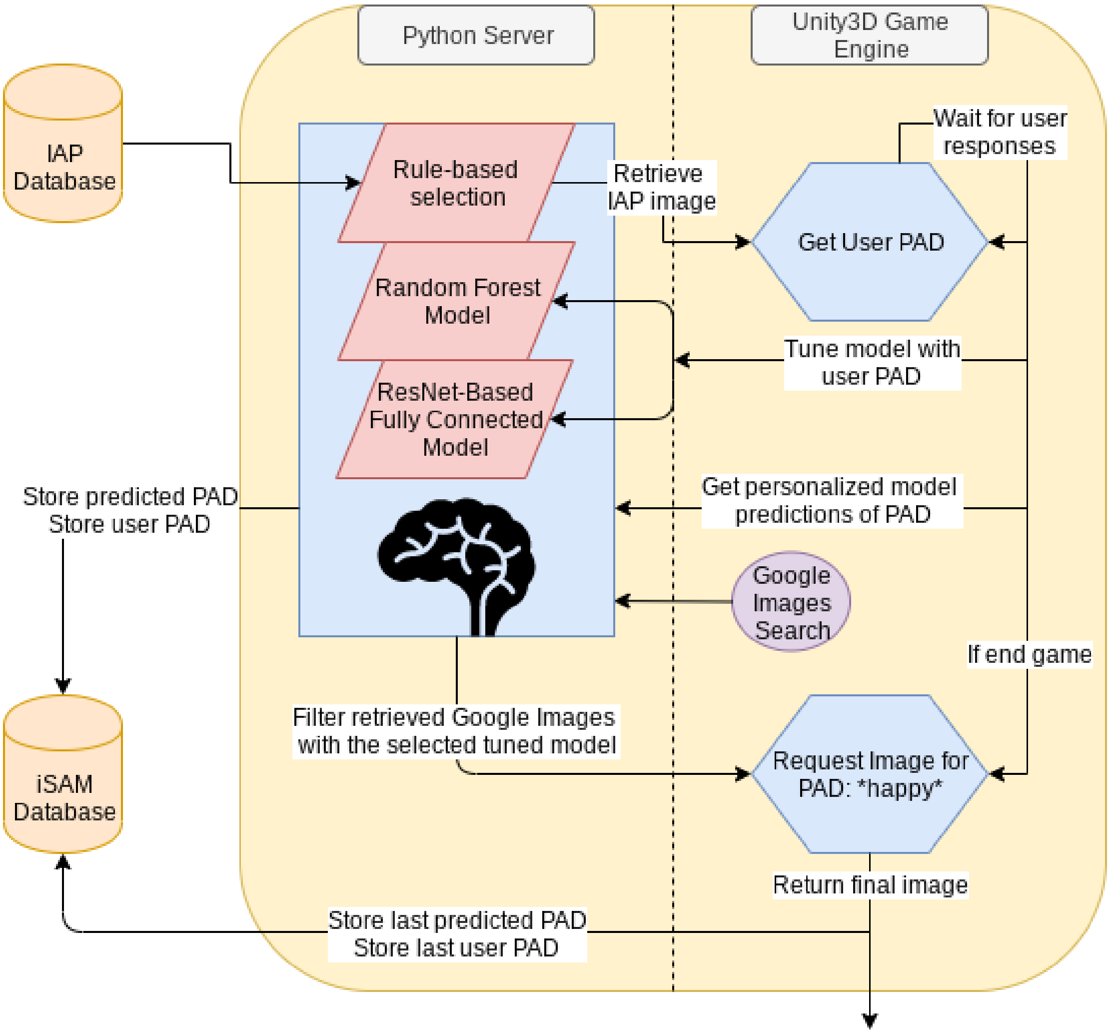


Figure 2.3: iSAM logical flow

Emotion Type	P	A	D
Happy	9	5	5
Surprise	9	9	5
Sadness	1	1	9
Disgust	1	1	1
Fear	1	9	9
Anger	5	9	1
Neutral	5	5	5

Table 2.1: PAD inference assumptions for the iSAM extreme emotional states. The scale in Self-Assessment-Manikin is from 1 to 9 [3, 4, 5, 6]. The assumptions above were used to retrieve IAPs references with a 1f delta random range. IAPs Images are non-repeating. Final curated images were produced by evaluating user CSV files to search for matching PAD assumptions. The predicted emotional state of the subject and the PAD assumptions were fed into the AI model to get the best matching "happy" image. (P → 1 unpleasant to 9 = pleasant, A → 1 calm to 9 excited, and D → 1 independent to 9 dependent [7])

female participants, one from male participants, and one from all participants combined [8]. For our pilot test, we used the metadata version from all participants – because our user pool was so small, it did not make sense to segment by gender.

Two control methods were designed for both the environment and player input. Environment control is enabled through an IAPs picture frame in the Unity Scene, as shown in Figure 2.4, and handles the updates and communication to the IAP database. We implemented the frame through the Unity3D's *Texture2D* class attached to a game object allowing us to switch images in and out of the picture frame. Secondly, user input control was curated through a 3DUI representing SAM. With simplicity in mind, we used trigger colliders on the VR controller to enable button presses on a world level canvas. Thus, iSAM can be played by any VR device that has a controller, and other affective survey formats can be easily translated into the game environment by directly importing sprites. The VR controller is represented by a rectangular pointer, as seen with the yellow sphere in Figure 2.2 near the SAM interface. A button click is done by passing the VR controller through the desired PAD value. Once a valid PAD value is given to every dimension, the “Confirm SAM Input” button becomes available for the user to store their response. Subsequently, this confirmation builds a string with a global timestamp that holds response time and PAD value to be stored and communicated in CSV format by the python server.

As user PAD responses come in, they are sent to the iSAM database to be stored. The user PAD responses are also sent to the python server to fine-tune the affective model component (shown in Figure 2.3 as the box with a brain icon). We designed three affective computational models – a base model that is simply the emotional data from the IAP database, a random forest model trained on the base model, and a fully-connected model trained on the base model using ResNet50 features. For the base-model, PAD prediction is possible for images already in the IAPs database, so the selection of the final image is only made by inferring an appropriate image category. The random forest and fully-connected models can be trained on updated data sets and can output PAD predictions for any image. In the interest of time, only the fully-connected model has all server endpoints implemented. After fine-tuning, predicted PAD using the updated Machine Learning (ML) model was also sent to the iSAM database to be stored.

2.4 User Evaluation

The iSAM prototype was evaluated with an initial set of four users. We set out to see if iSAM can accurately account for user emotion and if iSAM had any critical design flaws or improvement needs. To accomplish this, we employed a mixed-method approach of task-based evaluation and system log-file analysis. Four university students (1F, 3M, aged 22 to 27) were recruited to play-test iSAM at a home office nearby

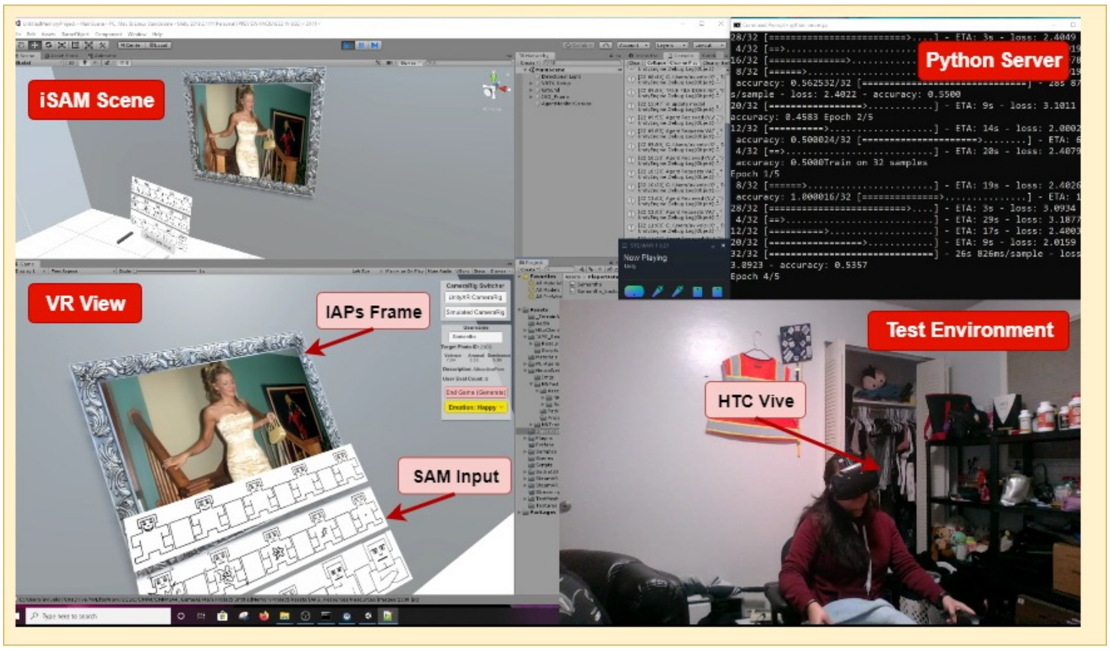
Inference Type	Pleasure	Arousal	Dominance
IAPs Database	1.0776 (E 0.2)	0.1827 (E 0.1)	0.1813 (E 0.2)
Static AI Model	2.5784 (E 0.1)	1.8947 (E 0.14)	0.1315 (E 0.05)
Tuned AI Model	1.9251 (E 0.08)	1.8275 (E 0.15)	0.8520 (E 0.06)
Model Increase	32%	1%	-86%
End Static Model	**3.5634 (E 0.9)**	**1.9167 (E 0.1)**	**1.2302 (E 0.6)**
End Tuned Model	**2.9651 (E 0.9)**	**1.5555 (E 0.2)**	**0.2706 (E 0.5)**
End Model Increase	**17%**	**21%**	**354%**

Table 2.2: Median results on user inference difference PAD responses. The upper half indicates IAPs training data images (N=40); the lower half indicates the final AI curated Google image after training data PAD responses (N=4). E indicates a standard error; the scale is in Self-Assessment-Manikin from 1 to 9. Note a lower value corresponds to closer accuracy to user PAD response.

the university campus. Play-testing sessions lasted from approximately 15-20 minutes, where users were introduced to iSAM by two research evaluators. Users were instructed that they would be helping SAM recover its lost memories by appraising memory gallery photos in virtual reality. The users wore an HTC Vive iVR HMD and were tasked with appraising ten images from the IAPs data-set by utilizing SAM. After the ten images are appraised, the AI model generates a final image based on the subset of the closest matching user reported PAD photos. This image is meant to be closest to the emotion goal state, shown in Table 2.1, where the user then evaluates the final AI curated image. Figure 2.4 demonstrates one of these four users play-testing iSAM with the described protocol above. After play-testing, we asked the users if they had any additional comments and how they felt about the final image. We then synchronized all recorded data in the iSAM database post hoc for statistical analysis.

2.5 Results and Discussion

Four users playtesting iSAM to each assessed ten training images from IAPs for fine-tuning, in addition to evaluating a final AI curated image. This results in a



```
28/32 [=========================>....] - ETA: 3s - loss: 2.4049
 4/32 [==>...........................
16/32 [==============>...............
 8/32 [======>.......................
 accuracy: 0.562532/32 [==============================] - 28s 87
s/sample - loss: 2.4022 - accuracy: 0.5500
20/32 [=================>............] - ETA: 9s - loss: 3.1011
accuracy: 0.4583 Epoch 2/5
12/32 [==========>...................] - ETA: 14s - loss: 2.0002
 accuracy: 0.500024/32 [=====================>........] - ETA: 6
 4/32 [==>...........................] - ETA: 20s - loss: 2.4079
 accuracy: 0.5000Train on 32 samples
Epoch 1/5
 8/32 [======>.......................] - ETA: 19s - loss: 2.4026
 accuracy: 1.000016/32 [==============>...............] - ETA: 1
28/32 [=========================>....] - ETA: 3s - loss: 3.0934
 4/32 [==>...........................] - ETA: 29s - loss: 3.1877
12/32 [==========>...................] - ETA: 17s - loss: 2.4003
20/32 [=================>............] - ETA: 9s - loss: 2.0159
32/32 [==============================] - 26s 826ms/sample - loss
3.0923 - accuracy: 0.5357
Epoch 4/5
```

Figure 2.4: iSAM Gameplay

total of 40 training images and four final evaluation images. Users said they found iSAM to be interesting and found the final image to feel happy – these final images produced by each user can be found in Figure 2.5. Table 2.2 reports the median training image offset between the user's PAD scores for multiple inference types. From the IAPs Database, mean inference was reasonably accurate in predicting user response (best 0.18 offset, worst 1.07 during training images). This indicates consistency with Lang et al's evaluation of IAPs with SAM [8, 289]. In terms of training, the AI Model's tuned performance demonstrated that it learned from the user. The difference in model tuning compared to its static training data showed significant differences after a training session for the final curated image, as demonstrated by the 32%, 1%, and -86% changes from pleasure, arousal, and dominance respectively.

The final generated images were found to be successful in producing a PAD response indicating happiness based on the majority of verbal and PAD responses. The lower half of Table 2.2 reports the median final generated images (as shown in Figure 2.5) and their offset between the user's PAD scores for multiple inference types. The tuned AI model demonstrated a significant percent change from the tuned model on user responses compared to the statically trained model on the IAPs database. These values all displayed a drop in SAM offset, indicating that the tuned model became far more accurate from the users' training responses with significant gains in arousal and dominance predictions.

While the initial results of iSAM were promising, we must consider some limitations. More users must test iSAM to fully determine its success in predicting and adapting to PAD emotional response. Additionally, more emotional groups beyond happiness should be considered in these playtest sessions. In terms of the AI model, PAD prediction accuracy and tuning can be made more optimal with more training time and with IAPs baseline data enhancement techniques. These limitations are being considered for future studies with our pilot data in mind.

2.6 Conclusion

Through the iSAM prototype, we presented a novel playable experience that employed AI and immersive virtual environments to learn from and adapt to a user's PAD emotional model. We demonstrated a pipeline to enable both users and AI analysis of the International Affective Picture System through a Virtual Reality interface that transported users into a "mind museum" to help SAM recover its lost memories. Our initial play-testing indicated that the AI model was able to improve its PAD emotional prediction for the majority of users through ten training photos from the IAPs and a final curated photo from Google images that was intelligently selected based on the

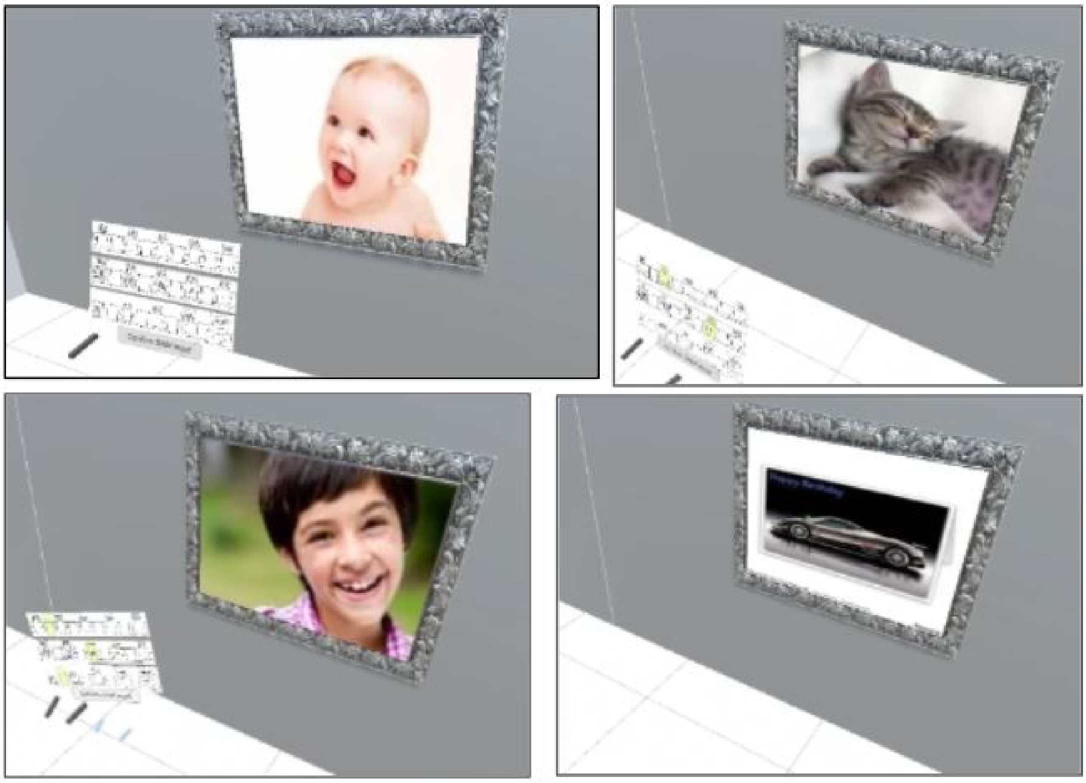


Figure 2.5: iSAM's final images produced for each of the four users. These photos indicate optimal PAD values for subjects of babies, children, kittens, and race cars that were chosen by inferring PAD for each individual user from Google image results using AI.

user's PAD responses. Subsequently, this work suggests that it may be possible to bridge runtime emotional models into a virtual environment, which may have substantial implications for the serious games community or any researchers interested in translating emotion models into immersive virtual environments through game engines and AI. For the future, we hope to explore this experience with multi-modal biofeedback to help influence the AI actions and user evaluation. While we feel that this is a step in the right direction, more work must clearly be done to verify the efficacy of iSAM and AI informed experiences that work off of the Pleasure-Arousal-Dominance emotional model. Subsequently, there are far more galleries of the mind to explore and memories to recover in the road ahead.

2.7 Acknowledgments

The authors would like to thank the following UC Santa Cruz affiliates for their support and guidance on this project: Conrad Esch, Samantha Conde, Michael Mateas, and Sri Kurniawan.

Chapter 3

Understanding Emotional Expression with Haptic Feedback Vest Patterns and Immersive Virtual Reality

3.1 Summary

Haptic feedback vests afford a unique opportunity to enhance a user's emotional engagement within the virtual world. In this paper, we present a two-stage study on user experience towards understanding emotional expression through utilizing the Pleasure-Arousal-Dominance model and the International Affective Picture System. We examine an authoring survey with 40 young adults, where users contextualized five emotion groups and designed patterns to express the feelings associated with each stimulus. Our resulting content analysis suggests design themes on a body-mind situation axis and an internal-external location axis. We found that vibrotactile actuation extends user emotion through phenomena which we call scene emulation, body function emulation, emotional resemblance, and emotional reflection. Lastly, we pilot these findings

through an immersive virtual reality experience.

3.2 Introduction

Immersive Virtual Reality (iVR) can be a powerful tool for understanding human behavior, where head-mounted display systems (HMDs) enable full-body movement and 360°viewing of the virtual world. In 2019, seven million commercial HMDs were sold and were projected to reach 30 million sales per year by 2023 [35, 52]. Such immersive input modalities with commercially available components may help us interpret user performance in serious games for player compliance, accessibility, and data throughput [15, 16, 14]. Researchers have reported considerable success in using virtual environments with serious games that explore psychological and physiological applications in various case studies and theories [290, 69, 19]. The detachment from reality driven by immersion with an iVR HMD can reduce discomfort for a user, even as far as minimizing pain when compared to clinical analgesic treatments [61, 62]. Immersion can be enhanced by improving graphics, multi-modality, user interaction, and emotional engagement [69]. Moreover, contextualized stimuli allow the user to feel present and emotionally engaged in the virtual world [64, 70, 71]. Modern-day game engines such as Unity3D provide flexible methods of creating immersive media in any operating system, enabling visualizations and interactions for construction, entertainment, government, and healthcare [291]. In addition to audio-visual experiences, haptic feedback input devices, such as vibrotactile vests, are becoming commercially available. In this study, we evaluate how haptic feedback vest patterns can extend emotional expression and boost feeling of presence for users.

3.2.1 Related Work

Emotion can be quantified through the Pleasure-Arousal-Dominance (PAD) emotional model, which uses three independent bipolar dimensions to quantify human emotion: pleasant-unpleasant (P), aroused-unaroused (A), and dominant-submissive (D) [285, 286, 292]. One way to obtain the PAD state of a user is the Self-Assessment Manikin (SAM), a visual survey that enables non-verbal, language-free, and quick retrieval of PAD for a given stimulus [287]. Applications of SAM include the 1997 CSEA (University of Florida's Center for the Study of Emotion and Attention) affective databases for pictures, audio, and words [8]. Each of the CSEA affective databases includes usability guidelines and contains a corpus of pre-tested stimuli with mappings of media-to-affective values through PAD [7, 8]. As discussed in our pilot study section, we utilize the picture corpus of the CSEA affective database, also known as the International Affective Picture system (IAPs).

Previous studies examined how haptic feedback can magnify and communicate pleasure and arousal and inform haptic feedback design. Bailenson et al. shows that users are able to both recognize and communicate emotions through a hand-based force-feedback haptic joystick [111]. Mazzoni et al. found haptic gloves can effectively map emotions from music to convey pleasure and arousal [112]. Salminen et al. investigated the patterns of a friction-based haptic system by horizontally rotating fingertip stimulators for pleasure, arousal, approachability, and dominance for hundreds of different stimuli pairs [114]. Fingertip actuation indicated that direction and frequency change of the haptic stimulation led to significantly different emotional information. And Miri et al. examined the design and evaluation of vibrotactile actuation patterns for breath pacing to reduce user anxiety and found that frequency, position, and personalization are critical aspects of haptic interventions for social-emotional applications [116].

Additionally, some studies within the past three years have started to examine haptic vests for communication of emotions in iVR. Goedschalk et al. explored the commercially available KorFX vest to augment aggressive avatars, but found an insignificant difference between the haptic and non-haptic conditions [121]. Krogmeier et al. investigated how a bHaptics Tactisuit vest can influence arousal, presence, and embodiment in iVR through a virtual avatar "bump", and found significantly more embodiment and arousal with full vest actuation compared to no actuation [122]. However, this study only examined a particular pattern and one stimulus set.

3.2.2 Study Goals and Contribution

From our literature review, we found little work on the investigation of how varying vest patterns influence emotional response and expression, especially in iVR environments. To examine how these vibrotactile input devices can emulate a user's emotional expression, our study uses the bHaptics Tactisuit Vest, the PAD emotion model, IAPs, the HTC Vive HMD, and the Unity3D Game Engine. The bHaptics vest is a commercially available haptic feedback vest that provides a combined 40 actuation points for the front and back torso, a mobile and web design app to experiment with haptic vest patterns in real-time, and an SDK for the Unity3D Game Engine [293, 291]. Specifically, the goals of this study are to pursue the following:

1. To establish an understanding of how haptic vest patterns can extend emotion;
2. To determine design themes for haptic vest patterns in contextualizing emotional response using IAPs and PAD;
3. And to pilot how such themes with haptic vest patterns can be applied in iVR

and influence emotional response, by extending a previous iVR experience that was designed to capture emotion through an immersive Self-Assessment Manakin [18].

To the best of our knowledge, this study is one of the first to establish emotional pattern design themes for upper body haptic vests. We explore these themes through two stages: examining the authoring of vest patterns for a variety of emotion groups with 40 young adults, and piloting the reaction of vest patterns within an immersive virtual environment with 4 young adults.We hope our analysis of haptic patterns may lead to insights towards more emotionally relevant input device techniques for serious game developers, researchers, and evaluators.

3.3 Design Pattern Authoring Study

We created a survey to examine how haptic vest patterns can extend and contextualize a user's emotional state for our pilot study. We targeted the IAPs photo stimuli to explore a combination of low, medium, and high pleasure-arousal (P,A) images. The IAPs provides a labeled database of ten thousand previously examined images using SAM with over 300 participants from 1997 [8, 7]. To examine different emotional ranges, we decided to vary the pleasure and arousal of various emotional states and kept dominance as neutral as possible. From this database, we are able to sort the images to find the closest matches for the following emotion sets: Happy (P,A → 9,5), Surprise (P,A → 9,9), Sad (P,A → 1,1), Neutral (P,A → 5,5), and Anger (P,A → 5,9). Inference assumptions were made for pleasure-arousal values of each emotion group after reviewing the IAPs and SAM technical reports [7, 8, 18]. We chose eight of the top images for each emotional group from the IAPs, leading to 40 images for the five emotion sets. A picture selection component of the survey, as shown in Figure 3.1, is utilized for users to pick which picture best matches the feeling of the emotion set.

To align the user baseline for explaining emotions, we employed SAM and the pleasure-arousal dimensions as a mechanic of our survey [7, 286]. We asked our users to perform a pre and post-test SAM to examine emotion change and followed the IAPs technical manual explanations for each emotion dimension, as seen in Figure 3.1. This provided a medium for non-verbal expression of emotion through the pleasant-unpleasant and aroused-unaroused SAM, which we then used to query users on selecting images for each emotion set. Additionally, we could see how working with the bHaptics vest design process itself would vary pleasure and arousal and added two post-test questions on user comfort and self-reported confidence of pattern design.

For designing haptic patterns, we utilized the bHaptics Designer mobile app to enable survey participants to feel and test various self-designed haptic vest patterns. The app allows for runtime actuation of the 40 vibrotactile vest positions on the front and back torso through touchscreen controls. The user toggles different positions on either the front, backside, or both sides of the vest through a draw or point-click mode while also varying intensity (from 0 to 100%) and duration (in time) of a given pattern. We asked our users to design a pattern to “extend the feeling” of their selected picture through the haptic vest. Specifically, users were tasked with drawing self-selected patterns on the bHaptics Designer app. They recorded the final pattern on a virtual drawing sheet, as shown in Figures 3.1 and 3.2. Users were then required to explain the reasoning behind each designed pattern and how it related to the feeling of each photo.

Choosing the most relevant image from the emotion set of eight photos enabled us to contextualize the IAPs for the year 2020. Querying SAM pleasure-arousal enabled

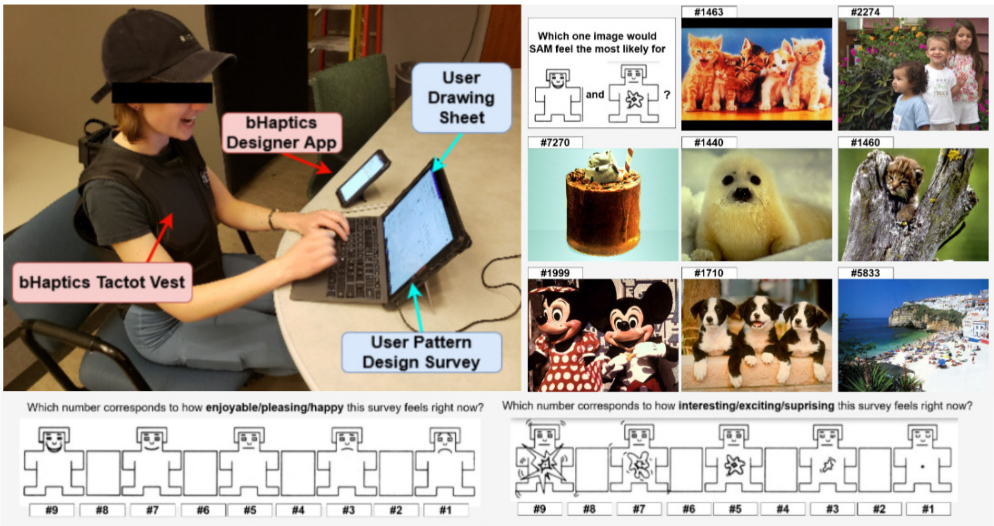


Figure 3.1: A user participates in the haptic vest pattern design survey (top left). The user chooses one top image from each of the five emotion sets (the happy emotion set can be seen at the top right of the figure), tests patterns with the bHaptics designer app, records patterns on a surface tablet, and writes in their interview response for their reasoning. Pleasure-Arousal is recorded Pre and Post-Test with a SAM Survey Questions (seen at the bottom of the figure).

us to examine how images and patterns are designed based on emotional relevance. Furthermore, the bHaptics designer app enabled us to explore the authoring process of the vest and how emotions extend to upper torso vibrotactile actuation. We proceeded to create a testing protocol after our picture stimuli, emotion model, and haptic vest pattern design mediums selected.

3.3.1 Survey Protocol

With our survey design sorted, we designed a multi-stage experimental protocol to see how participants express their emotions through vibrotactile vest actuation. The experimental equipment consisted of a touchscreen tablet (containing the survey, on-screen pen, and draw-in sheet for vest patterns), a smartphone (containing the bHaptics designer app), and the bHaptics vest as seen in Figure 3.1. Our user testing sessions consisted of the following protocol steps:

Preparation - A study administrator sanitized the study equipment, made sure all equipment was fully charged, and ran an actuation test with the bHaptics vest to ensure data communication quality.

Researcher Administrator Instruction - The administrator instructed the user to remain still and relax. The user was verbally instructed to follow the survey and read the IAPs SAM information carefully. The research administrator then explained the bHaptics vest through discussing the 40 vibrotactile actuation points and giving three pattern examples to the user: a mid-section line on the user's front torso, a circle on the user's back torso, a wave on the user's front and back torso (synchronous), and toggled points outside of draw mode with intensity gradually lowered from 100% to 0%. The users were then asked to re-create the same patterns using the app themselves before proceeding. Afterward, the research administrator instructed the user to proceed to the survey and remained in the room to answer any technical questions about the equipment. Administrators did not answer any questions about the pictures or subject of the pictures themselves. They would respond to tell the user there is no wrong or right answer, and they emphasize the picture and pattern are whatever they perceive it to be. This study equipment is illustrated in Figure 3.1.

Survey Introduction - The participant was required to read the IAPs technical manual briefing on SAM and the pictures. This information conveyed all answers would remain anonymous. The user was asked to answer how they see each independent picture honestly and was given instruction on how to interpret pleasure and arousal from Figure 3.1 by Lang et al.'s explanation of SAM [8]. The user then filled out a SAM for pleasure-arousal and responded how comfortable they were in expressing different haptic feedback patterns.

Picture Set Evaluation (repeated five times for each emotion group) - Users evaluated a picture set of eight images by the target SAM of the emotion set (for example, Happy is a pleasure of nine and an arousal of five, as shown in Figure 3.1). After choosing which picture was the most relevant to the emotion goal, the user then designed a haptic feedback vest pattern using the mobile application. Once ready, the user recorded the pattern by drawing it onto the tablet. The user was then asked to describe the pattern they designed and how it related to the feeling of the picture. This stage was repeated until the user completed all emotion sets.

Survey Conclusion - After evaluating each emotional picture set, the user then filled out a second SAM for pleasure-arousal and responded how meaningful they thought their patterns were. The user then responded to how comfortable the haptic feedback vest was, and had the option to share additional comments with the research administrator.

Figure 3.2: Haptic vest pattern themes from the authoring study. Four exclusive pattern themes are identified on a two-axis domain focused on position (y-axis) and design reasoning (x-axis). Participant drawings reflecting each theme as an example can be seen with the outer boxes.

3.3.2 Pattern Authoring Survey Results

We successfully tested the design authoring survey with 40 participants, where sessions lasted about 25-45 minutes. All users reported having no prior experience in using haptic feedback vests. 29 of 40 participants reported demographics information, with a mean age of 27 years old (6.4 standard deviations), and 35% female, 3% non-binary, and 62% male self-reported gender identity. A total of 200 drawings and written interviews were recorded (5 emotion sets x 40 users). 85% of users self-reported they thought their patterns were meaningful – the rest were undecided. 80% of users self-reported the vest was comfortable, 7.5% said it was undecided, 7.5% said it was uncomfortable, and the rest gave other comments, i.e. *"Feeling like batman"*. The users that found the vest to be uncomfortable shared *"the front gives a bit of a choking sensation"* (P1), and *"the vest is like a life jacket... I wish it was like a sweatshirt"* (P26). From examining the emotion changes induced by the vest during the survey, a significant increase in self-reported arousal and an insignificant increase of pleasure was found, as shown in Table 3.1. These results are in line with Krogmeier et al., where haptics induce significantly more arousal [122]. Users echoed these results in their additional comments; for example, P5 shared they *"never designed haptic feedback patterns before, this was an exciting experience and felt powerful."*

To explore how the IAPs have changed over time and determine which pictures are the most relevant for our user demographics, we organized the top voted pictures in Table 3.2. Based upon user votes distribution, the Happy set was found to be the most distributed, and the Neutral set was the least distributed. A lower distribution was seen in the Neutral, Sad, and Surprising image sets, which may differ from the baseline IAPs data. In other words, some images considered to be Neutral, Sad, or Surprising in the 1997 data set are not as emotionally relevant for our users in 2020. For example, a picture of a battleship had a neutral IAPs PA score from 1997, and only one out of our forty users in 2020 resonated with this picture. Conversely, the Happy picture set was very distributed where users were choosing between photos of animals, beaches, and kids. These results may indicate happy images tend to transfer well over time and sociodemographic factors, and neutral and sad images do not.

	Pre-Test		Post-Test		**Difference**	
Metrics	P	A	P	A	**P** ^{ns}	**A** ***
Mean	6.75	4.13	6.97	5.93	**0.25**	**1.80**
(std)	(1.44)	(2.10)	(1.71)	(2.01)	**(1.59)**	**(2.68)**
Median	7.00	4.00	7.00	6.00	0.00	2.00
(QR)	(1.25)	(2.50)	(2.00)	(2.00)	(1.00)	(4.00)

Table 3.1: Self-reported user pleasure(P)-arousal(A) from the haptic vest pattern design survey (N=40). Change induced by survey indicated a significant increase in self-reported arousal ($p \leq 0.001$) and a insignificant increase of pleasure ($p > 0.05$) from t-tests.

Emotion	ID	Votes	Subject	Category	IAPs P	IAPs A
Happy	1463	8	Kittens	Animal	7.45	4.79
Happy	1440	6	Seal	Animal	8.19	4.61
Happy	1710	6	Puppies	Animal	8.34	5.41
Surprise	8492	10	Rollercoaster	People	7.21	7.32
Surprise	8185	9	Skydivers	People	7.57	7.27
Surprise	8370	5	Rafting	People	7.77	6.73
Sad	9220	11	Cemetery	People	2.06	4.00
Sad	9001	8	Cemetery	Scene	3.10	3.67
Sad	2490	7	Man	People	3.32	3.95
Neutral	1908	24	Jellyfish	Animal	5.28	4.88
Neutral	6000	6	Prison	Scene	4.04	4.91
Neutral	1230	5	Spider	Animal	4.09	4.85
Anger	8475	9	Biking	People	4.85	6.52
Anger	5950	7	Lightning	Scene	5.99	6.79
Anger	5920	6	Volcano	Scene	5.16	6.23

Table 3.2: Top 3 images most chosen by users for each emotion group. IAPs indicates baseline pleasure-arousal reported by Lang et al [8]. Votes (N=40) indicate distribution of user choices between top three images.

Pattern Theme	Scene Emulation	Emotional Resemblance	Body Function Emulation	Emotional Reflection
Percent Agreement	95.0%	90.5%	96.5%	88.0%
Scott's Pi	0.865	0.763	0.906	0.661
Cohen's Kappa	0.865	0.763	0.906	0.662
Krippendorff's Alpha	0.865	0.764	0.906	0.662
N Agreements	190	181	193	176
N Disagreements	10	19	7	24
Happy Occurrences	6	**14**	6	**14**
Surprise Occurrences	**14**	7	10	9
Sad Occurrences	4	11	**17**	8
Neutral Occurrences	**20**	11	4	5
Anger Occurrences	7	9	11	**13**

Table 3.3: Thematic coding between two researchers of haptic patterns for inter-reliability of over 200 cases in Percent Agreement, Scott's Pi, Cohen's Kappa, and Kippendorff's Alpha [9]. The number of occurring pattern themes per emotion group is shown in the lower half of the table.

Emotional Group	Top Pattern Theme	McNemar Test	p-value
Happy	Emotional Resemblance	**5.085**	**0.024**
Surprise	Scene Emulation	0.778	0.378
Sad	Body Function Emulation	**8.022**	**0.005**
Neutral	Scene Emulation	1.000	0.317
Anger	Emotional Resemblance	**5.232**	**0.022**
Image Category	Top Pattern Theme	McNemar Test	p-value
People	Body Function Emulation	**27.272**	**<0.001**
Animal	Scene Emulation	3.462	0.063
Object	Body Function Emulation	**46.080**	**<0.001**
Scene	Emotional Reflection	0.643	0.423

Table 3.4: Top pattern design themes by emotional group and image category over 200 cases. Uses pattern theme coding from coder 1. See Table 3.3 for inter-reliability. Significant categories are in bold.

iVR User	Usability (SUS)	Presence (Slater)	P Diff	A Diff	Scene Emulation	Body Function	Emotional Resemblance
User 1	32.5	3.2	2.00	2.00	7.00	6.00	2.00
User 2	20.0	5.3	0.00	0.00	8.00	2.00	5.00
User 3	35.0	3.5	0.00	0.00	6.00	5.00	4.00
User 4	27.5	4.7	0.00	1.00	6.00	4.00	5.00
Mean	28.7	4.2	0.50	0.75	6.75	4.25	4.00
Std	6.6	1.0	1	0.95	0.95	1.71	1.41

Table 3.5: Results of four users from immersive experience testing. Usability is in SUS Scale where 0 is usable and 100 is unusable [10]. Presence is in Slater's Scale where 1 is not present in the virtual environment, 4 is somewhat present, and 7 is fully present [11]. P Diff and A Diff are the difference in pre and post test pleasure-arousal scores from the user. The last three columns are the vote counts for which themes the users preferred for the 15 photos during testing.

3.3.3 Design Themes

With our initial results in mind, we proceeded to apply a thematic analysis to determine patterns in emotion between users. Two researchers performed thematic coding for each user's responses to image selection, haptic pattern drawing, and written feedback. Four design pattern themes were identified, suggesting a two-axis design domain with an external-internal axis of pattern position and a body-mind axis of design reasoning (Figure 3.2).

Scene Emulation: The vest pattern emulates physical sensations of external phenomena. These patterns would often expand the photo's physical environment to mimic or better contextualize a subject. For example, with a picture of a rollercoaster, P10 shared their pattern is a *"bumpy ride, top and bottom always on for front and back while the middle ones oscillate out of phase with each other, it has high intensity and movements like a rollercoaster."*

Body Function Emulation: The vest pattern emulates body reaction, such as a heartbeat or breathing. These patterns would often expand upon the bodily effects of the photo for how a body would act under a given scenario – a mimicking of body parts expressed through the haptic vest. For example, with a picture of ice cream, P1 reported their pattern to be a *"stomach rumbling; a vibration pattern across the abdomen region; I would look at the dessert and feel hungry and happy.*

Emotional Resemblance: The vest pattern is designed to change or suggest an emotional state. These patterns were often intended to make the receiver feel a certain way that did not mimic. For example with a picture of kids, P3 reported *"kids are calming, playful, warm and comforting. The back massage [their designed vest pattern] is similar to that feeling."*

Emotional Reflection: The vest pattern mimics the user's reaction towards a certain photo. The user often either drew or created a pattern based on an emotional reaction of a photo's subject, flow, or location with no direct intention for emotion change. For example, with a picture of a frowning old man, P35 designed their pattern *"just to make sadness sink in."*

These themes and each respective drawing for the described examples are illustrated in Figure 3.2. The two researchers assigned an exclusive thematic code to all 200 patterns with the four design pattern themes determined by comparing drawings, interviews, and picture choice. A series of inter-reliability tests were performed, as shown in Table 3.3. The inter-reliability metrics indicate substantial agreement and almost perfect agreement for all pattern themes between the two coders in terms of observer agreement for categorical data [9]. Independent themes were organized by emotion set as shown in Table 3.3.

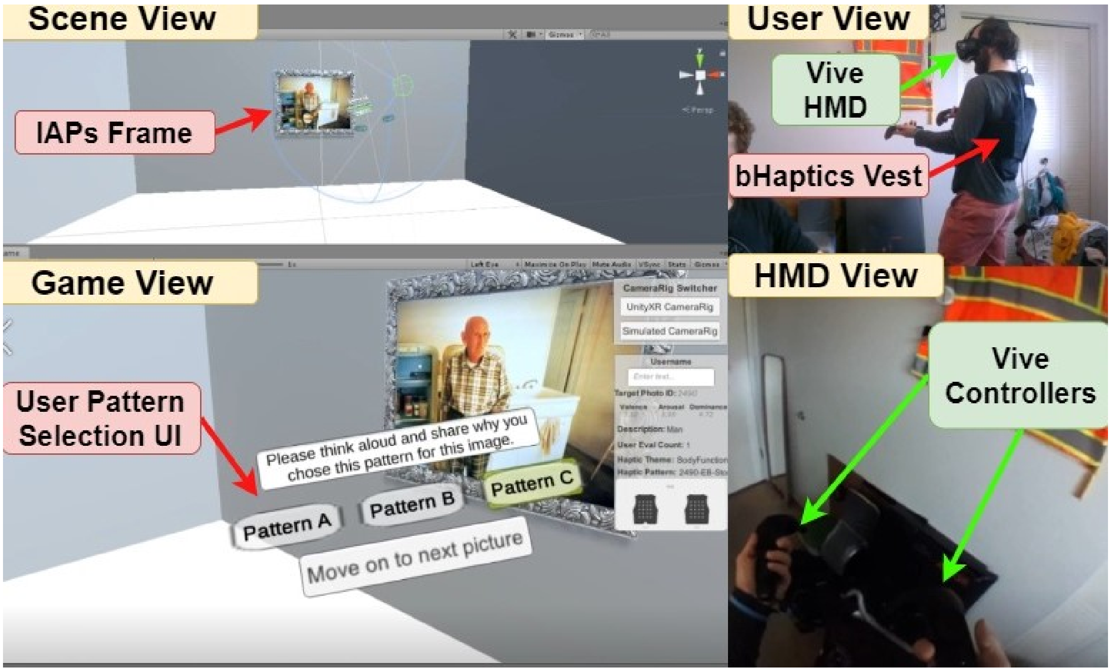
3.4 Flipped Pilot study in Immersive Virtual Reality

Our pilot study found four common themes, as shown in Figure 3.2, which occurred during haptic feedback pattern vest design. Users generally felt and expressed their reactions of photos through a body-mind rationale and an internal-external pattern positioning. We desired to explore the applications of these themes and approaches from a reactionary perspective rather than authoring. As a result, we created an iVR experience to "flip" the survey-based pilot study, expanding upon prior work on establishing an immersive Self-Assessment Manakin [18].

We designed our iVR system to automate some of our data collection processes while also counterbalancing stimuli on the fly. In this experience, a participant enters a virtual museum with the HTC Vive HMD and evaluates bHaptics vest patterns that correspond to the feeling of the presented IAPs photos. Through the top identified pilot study images in Table 3.2, we designed a set of haptic patterns for Scene Emulation, Body Function Emulation, and Emotional Resemblance. Emotional Reflection was excluded as it was the least haptic feedback oriented and required the developer to account for individual user memories and immediate reactions of the haptic designer. For the three themes included, patterns were designed by researchers through reflecting on the most re-occurring characteristics of participant drawings per photo subject and underlying author intentions. A total of 15 pictures (the top three voted photos from the pilot study of the five emotion sets) are presented to the user with 45 haptic patterns (three per each picture for Scene Emulation, Body Function Emulation, and Emotional Resemblance themes). Stimuli are automatically counterbalanced through the order of the emotion sets and each picture within every set to reduce order effect. The system records user response time, pre & post pleasure-arousal, vest pattern preference, and think-aloud audio interviews from users.

3.4.1 Preliminary Evaluation

Four participants from UC Santa Cruz were recruited with the same methods as the pilot study to examine the iVR experience. Participants were university students with a mean age of 25 (+/- 2.1) years old with one identifying as female and three identifying as male within the testing cohort. All of these users reported having some experience with using iVR and understood haptic feedback patterns from a survey screening. We expanded upon our pilot study protocol. Participants entered a virtual museum with the HTC Vive HMD and evaluated which bHaptics vest patterns best



Please think aloud and share why you chose this pattern for this image.

Figure 3.3: iSAM iVR experience for testing haptic vest patterns.

correspond to the feeling of the presented photos. Users were tasked with identifying the most emotionally relevant pattern per each photo they reviewed. After identifying their preferred pattern, users thought aloud to explain their reasoning, and interview audio was recorded and stored through the Unity3D Game Engine. At the end of testing, the iVR HMD was removed, and users filled out a System Usability Scale (SUS) survey from Brooke et al. [10] and a Slater Presence survey [11].

3.4.2 Immersive Experience Results

Each of the four users completed the iVR experience to evaluate the 45 vest patterns for the 15 photos. Sessions lasted approximately 20-30 minutes and were conducted at a home office. Pre and Post Pleasure-Arousal tests indicated users had a small increase in both emotion metrics similar to the authoring-based pilot study. Participants reported they found the iVR experience and the vest to be highly usable, as shown in Table 3.5. All four users found the system was quick to learn and not too complicated, and felt confident in controlling the haptic pattern selection. The users suggested improvement areas, including moving the confirm button placement above the pattern buttons and making the timing in the protocol quicker (Figure 3.3). For presence, about half the users felt above neutral presence, and half felt below neutral presence, as shown in Table 3.5. Users who reported a higher presence had a higher preference count for the Emotional Resemblance theme.

3.5 Discussion

We gathered some insights on how affective databases like the IAPs adapts over time and geography. In our pattern design survey, the Happy emotion group appears to adapt well, with users making various choices for the picture that best matched the emotion. In contrast, the Neutral, Sad, and Surprise emotion groups had users overwhelmingly choose particular pictures as best matches. For example, the Surprise group contained many pictures of people pursuing winter sports and snowfall activities that are not so applicable to our west coast user population. Having our users determine which photos were the most emotionally relevant for them also allowed us to ensure our immersive experience stimuli were contextual to our target users.

In examining the effects of the haptic feedback vest, we found the torso can provide an area for rich user expression and emotional reflection. Our design themes helped map user reflection through an intention-based body-mind axis and a location-based internal-external axis. Users would approach photos by psycho-physical mimicking, reflecting upon an immediate feeling, or a desired emotion to better contextualize their response. Depending on the emotion group, different users would generally approach patterns through a single pattern theme or vary to one or two patterns throughout their testing session. Overall, the pilot study results suggest haptic feedback on the torso is useful for improving users' emotional engagement.

For example, users shared some of the following perspectives: *"I feel that people think with their entire bodies, even if in a subtle way. So, the haptics being that much more of an enhancement seems like a really cool tool"* (P5), *"This haptic vest reveals something that I did not know before, e.g., I do not like any forms of vibration on my chest, and most of my emotions are related to my digestion system. That means that I feel more connection between emotion and tummy instead of my back"* (P37), and *"It's [expressing patterns with the vest is] kind of like the physical equivalent of the inkblot test"* (P3). As the consumer market provides wider availability for more immersive equipment beyond audiovisual devices, such as haptic vests, these findings may provide exciting future research opportunities. Other researchers may be able to expand upon these pattern themes to personalize haptic feedback for emotional engagement, and personalization is considered an essential element of using haptics for healthcare interventions [116].

To understand pattern themes, we designed and ran an evaluation of how vest patterns can be applied to iVR. We found users with a higher self-reported presence tended to prefer Emotional Resemblance over physical mimicking patterns more than users with a lower self-reported presence. This is particularly exciting as prior research suggests emotional engagement leads to a higher perceived presence [69]. The experience itself was found to be highly usable, and the data automation enabled through the Unity3D game engine helped optimize the user analysis process. However, more work is needed to evaluate the immersive virtual experience and must engage a larger sample size of non-experts of iVR as well as haptic pattern experts. We believe this initial work is a step in the right direction and hope it may inform future researchers interested in using haptic vest patterns for experiences centered around emotion.

3.6 Conclusion

Understanding how haptic feedback vests can extend emotional expression and reflection is an emerging topic that is under-explored. This paper demonstrated a methodology for understanding how haptic vest patterns can extend human emotion through a two-stage study on both designing haptic vest patterns and evaluating patterns in iVR. We explored how to contextualize the emotions of happy, surprised, sad, neutral, and angry for a group of 40 users through utilizing the International Affective Picture System. By coding user designs, we found haptic vests patterns can be mapped for emotion by a body-mind situation axis and an internal-external location axis. Four themes emerged for expressing emotion in haptic vest patterns: Scene Emulation, Body Function Emulation, Emotional Resemblance, and Emotional Reflection. To apply our findings, we illustrated a design and evaluation of an iVR experience that tasked four users on selecting which vest pattern theme felt the most relevant to them. Our findings suggested users who were more emotionally oriented tended to self-report a higher presence and usability level. This work may indicate haptic vest patterns for emotion provide rich insights for users and designers, and emotionally oriented patterns may lead to higher engagement within the virtual world.

3.7 Acknowledgements

We thank Professor Katherine Isbister of the UC Santa Cruz Computational Media Department for her advice during this study, and the many users who volunteered to participate in our evaluation.

Part IV

EXPLORING IMMERSIVE VIRTUAL REALITY EXPERIENCES FOR PHYSICAL INTELLIGENCE

Chapter 1

Introduction

This investigates how immersive virtual reality experiences can be personalized for greater physical intelligence from [16, 228]. We demonstrate how deep reinforcement learning can be utilized to create visually assistive agents in exercise games through Proximal Policy Optimization and General Adversarial Imitation Learning. We also examine the design and evaluation of optimized distributed gradient boosted algorithms for predicting key physical rehabilitation success metrics in immersive virtual reality exercise games with off-the-shelf headsets. This chapter concludes with considerations for utilizing immersive virtual reality with machine learning for telehealth and serious game analytics in physical rehabilitation.

Chapter 2

Deep Reinforcement Learning in Immersive Virtual Reality Exergame for Agent Movement Guidance

2.1 Summary

Immersive Virtual Reality applied to exercise games has a unique potential to both guide and motivate users in performing physical exercise. Advances in modern machine learning open up new opportunities for more significant intelligence in such games. To this end, we investigate the following research question: What if we could train a virtual robot arm to guide us through physical exercises, compete with us, and test out various double-jointed movements? This paper presents a new game mechanic driven by artificial intelligence to visually assist users in their movements through the Unity Game Engine, Unity Ml-Agents, and the HTC Vive Head-Mounted Display. We discuss how deep reinforcement learning through Proximal Policy Optimization and Generative Adversarial Imitation Learning can be applied to complete physical exercises

from the same immersive virtual reality game. We examine our mechanics with four users through protecting a virtual butterfly with an agent that visually helps users as a cooperative "ghost arm" and an independent competitor. Our results suggest that deep learning agents are effective at learning game exercises and may provide unique insights for users.

2.2 Introduction

Physical activity is an essential part of daily living, yet 48.3% of the 40 million older adults in the United States are classified as inactive [37, 38]. Inactivity leads to a decline of health with signification motor degradation: a loss of coordination, movement speed, gait, balance, muscle mass, and cognition [36, 37, 38]. The medical benefits of regular physical activity include weight loss and reduction in the risk of heart disease and certain cancers [39]. However, compliance in performing regular physical activity often lacks due to high costs, lack of motivation, lack of accessibility, and low education [38]. As a result, exercise is often perceived as a chore rather than a fun activity.

Immersive Virtual Reality (iVR) and the increasingly recent use of games for health and well-being have shown great promise in addressing these issues. The ability to create stimulating and re-configurable virtual worlds has been shown to improve exercise compliance, accessibility, and performance analysis [50, 42, 13]. Other studies have suggested that engaging in a virtual environment during treatment can distract from pain and discomfort while motivating the user to achieve their personal goals [294, 62]. Additional success has been reported in using virtual environments for a broad range of health interventions from a psychological and a physiological perspective [290, 69]. Some of the biggest challenges that these studies found were technological constraints such as cost, inaccurate motion capture, non-user friendly systems, and a lack of accessibility [77, 42, 33].

The past five years have seen explosive growth of iVR systems, stemming from a projected 200 million head-mounted displays systems sold on the consumer market since 2016 [52]. This mass adoption has been in part due to a decrease in hardware cost and a corresponding increase in usability. From these observations, we argue that the integration of iVR as a serious game for health can offer a cost-effective and more computationally adept option for exercise. These systems provide a method for conveying 6-DoF information (position and rotation), while also learning from user behavior and movement. While there has been a number of works in exploring iVR environments for physical exercise [13, 50, 69], we present our paper as an exploration of making these environments more physically intelligent through machine learning. Specifically, we leverage the integration of the Unity Game Engine, ML-Agents, Deep Reinforcement Learning, and a custom in-house iVR exercise game. Through these technologies, we examine how neural network agents can augment a playable experience where a virtual robot arm assists user exercise masked as a task of protecting butterflies from incoming projectiles.

2.2.1 Virtual Reality and Machine Learning

Virtual games provide controlled environments and simulations for a wide range of Artificial Intelligence and Machine Learning applications. Game AI has been extensively researched from mechanical control, behavior learning, player modeling, procedural content, and assisted gameplay [295]. Applying machine learning to the virtual game domain opens up a playground for researchers to find appropriate learning techniques and solve various reward-based tasks [296]. For example, Conde et al showcased reinforcement learning for behavioral animation of autonomous virtual agents in a town [297]. Huang et al demonstrated imitation learning through a 2D GUI to control a Matlab simulated robot in sorting objects [298]. Yeh et al explored Microsoft Kinect exercise with a Support Vector Machine (SVM) classifier for quantified balance performance [299]. Additionally, agent learning in an iVR environment may be especially advantageous for assistive applications.

The computational requirements and data-throughput of modern iVR systems can be leveraged to analyze therapeutic gamification [13, 300, 301], postural analysis [258], and accuracy for research data collections [259]. This is important because iVR systems must have accurate motion capture and low latency of a user's position and rotation from the physical world to reduce motion sickness [302]. As a result, iVR systems are becoming more powerful, immersive, accurate at capturing user behavior, and affordable to the average consumer [52].

Some researchers are recognizing the potential of utilizing machine learning and AI with iVR systems. Zhang et al explored an iVR environment for human demonstrated robot skill acquisition [303]. The authors describe a deep neural network policy to solve this problem for training teleoperation robotics and illustrate that mapping policies of learning using VR HMDs is challenging. Through utilizing an HTC Vive, PR2 Telepresence Robot, and a Primesense 3d camera, the authors successfully trained their neural network to control a robot by collecting user 6-DoF pose and color depth images of player movement. In terms of utilizing machine learning to support player movement, we found two recent studies through our literature review. Kastanis et al described a method of reinforcement learning for training virtual characters to guide participants to a location in an iVR environment [304]. The authors used presence theory to predict uncomfortable interpersonal distance for human players and successfully incentivized study participants to move away from trained virtual agents. And Rovira et al examined how reinforcement learning could be used to guide user movement in iVR through projecting a 6-DoF predictive path for user collision avoidance [305].

While several works have been explored in utilizing machine learning for games, and researchers have started looking at iVR as a medium for human-agent learning, there have been few works exploring agents for iVR exergaming. iVR exercises can provide a vehicle for real-time motion capture and inverse kinematics of player movement. Such data could enable the analysis of confounding postural issues, such as slouched backs and other movement biases, and could adapt the game in real-time to maximize exercise outcome. With these previous works in mind, we consider the following question: what if we could have a predictive model that could inform us of our movement trajectory in a virtual exercise game?

2.2.2 Study Goals and Contribution

The prior work discussed in this section has demonstrated that deep reinforcement learning can enable promising predictive models for system control and user behavior. Little work has been done in exploring machine learning from 6-DoF user exercise movement (or movement in general) for iVR experiences. Through this project, which we call "Illumination Butterfly (IB)," we aim to explore how deep reinforcement learning can inform iVR exergames in terms of user movements and game mechanics. Specifically, the goals of this study are to:

1. Examine Deep Reinforcement Learning for a Double-Jointed Virtual Arm to model

physical exercise movements through 6-DoF interaction with Immersive Virtual Environments.

2. Explore the capabilities of Generative Adversarial Imitation Learning (GAIL) and Proximal Policy Optimization (PPO) for learning in-game physical exercises.

3. Evaluate the trained agent for cooperative and competitive exercise applications between human users.

Our serious game explores neural network-driven 3DUI interaction techniques by using two emergent machine learning algorithms (GAIL and PPO) to see how a virtual robot arm can both cooperatively and competitively guide users in their movements. This project stems from previous iVR games designed through the interpretation of exercise theory and human anatomy. We expand our work from Elor et al's previous exploration into serious games for upper-extremity exercise movement: a multi-year interdisciplinary exploration between local healthcare professionals, roboticists, game developers, and disability learning centers at Santa Cruz, California [15, 13, 12, 227, 18]. Through leveraging machine learning, we hope to enable Project IB as a new computational experience to understand human exercise and robotic behavior via virtual butterfly. This project may be a step forward for other researchers interested in integrating “physical intelligence” via predictive models of user movement for other iVR exergames.

2.3 System Design

The system in this paper is based on “Project Butterfly” (PBF), a serious iVR game for exercise previously explored by Elor et al [15]. We heavily modified PBF to create a new gaming experience directed at AI guided upper extremity exercises. Our version of PBF was developed in the Unity 2019.2.18f1 Game Engine with SteamVR 2.0 and incorporates the HTC Vive Pro 2018 by Valve Corporation, a highly adopted commercial VR system that uses outside-in tracking through a constellation of "lighthouse" laser systems for pose collection in a 3D 4x4m space [52, 291, 257]. Vive has been verified in previous studies to analyze therapeutic gamification [13, 300, 301], postural analysis [258], and accuracy for research data collections [259].

The objective of the game is to protect a virtual butterfly from inclement weather and projectiles by covering the avatar with a translucent "bubble shield" using the HTC Vive Controller. Thus the player is required to follow the path of the butterfly with plus or minus 0.1 meters, which enables the dynamic control of pace and position for a prescribed exercise. The player is awarded a score point for every half second they successfully protect the butterfly, with both audio and haptic feedback to notify them that they were successful. By protecting the butterfly, the world around them changes - meadows become brighter, trees grow, and the rain slows down. Conversely, if the butterfly is not protected, no positive feedback occurs - the world does not change. The game can be tailored to each player's speed and range of motion through a dynamic evaluator interface. Previously, PBF was explored with post-stroke and older users to analyze the feasibility of the game with exo-skeletal assistance for two exercises [15] by Elor et al, but was not designed or tested for neural network guided upper extremity movements varying custom exercise movements as reported in this paper.

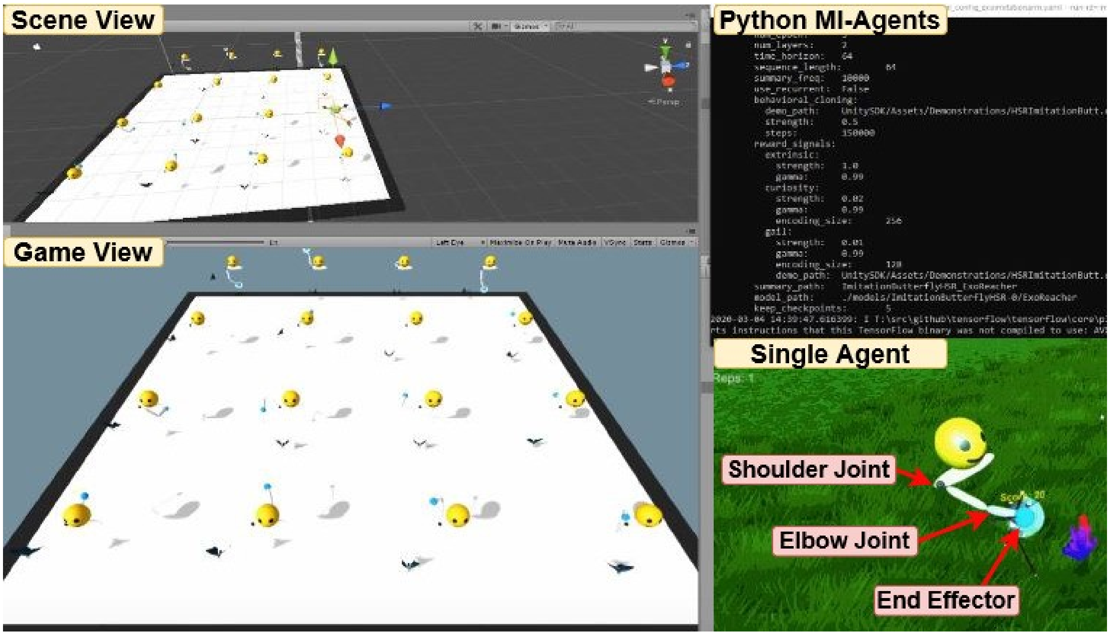
To explore the application of deep-learning agents for visually guided upper-limb exercise, we created a new modified version of PBF, which included the following changes from the previous version:

1. A modified “Reacher Agent,” a double-jointed arm controlled by predictive torque [306], was added into the player controller with the reward given when protecting a virtual butterfly.
2. A training scene for 16 parallel agents and three butterfly movements was created, as shown in Figure 2.1.
3. A “ghost arm” game mechanic was added for user visual guided movements with the original PBF game modes, and a “human vs agent” game mode was added for competitive analysis.

To the best of our knowledge, this study is one of the first to leverage an immersive VR HMD such as the HTC Vive with deep reinforcement learning to examine visually assisting agents for exergaming.

2.3.1 Machine Learning Environment and Agent Design

Project IB has been fully integrated with Unity ML-Agents, an open-source Unity plugin that enables games and simulations to serve as environments for training intelligent agents. The experimental plugin enables a python server to train agents in development environments through reinforcement learning, imitation learning, neuroevolution, and other emerging Tensorflow based algorithms [307, 308, 291]. We targeted upper-extremity torque and angular momentum as metrics to predict for our model. Having our AI model examine these metrics at the elbow and shoulder joints is advantageous. Torque is important as it used to describe the movement and force produced by the muscles surrounding the joint [309, 310, 311, 312]. Prior research has examined the torque of upper-body exercise for more in-depth injury assessment; for example,



```
    num_layers:     2
    time_horizon:   64
    sequence_length:        64
    summary_freq:   10000
    use_recurrent:  False
    behavioral_cloning:
      demo_path:    UnitySDK/Assets/Demonstrations/HSRImitationButt.
      strength:     0.5
      steps:        150000
    reward_signals:
      extrinsic:
        strength:   1.0
        gamma:      0.99
      curiosity:
        strength:   0.02
        gamma:      0.99
        encoding_size:      256
      gail:
        strength:   0.01
        gamma:      0.99
        encoding_size:      128
        demo_path:  UnitySDK/Assets/Demonstrations/HSRImitationButt.
    summary_path:   ImitationButterflyHSR_ExoReacher
    model_path:     ./models/ImitationButterflyHSR-0/ExoReacher
    keep_checkpoints:       5
2020-03-04 14:39:47.616399: I T:\src\github\tensorflow\tensorflow\core\p
rts instructions that this TensorFlow binary was not compiled to use: AV
```

Figure 2.1: Project IB Training Scene and AI Agents. Agents act as a double-jointed virtual arm with observation on the shoulder, elbow, and end effector joints. Sixteen agents were set up in parallel to train through the python ml-agents library with an action space of +/- 1.0 for actuating pitch and roll torques on the elbow and shoulder joints, respectively. A reward of +0.01 is given to the agent per every frame the end effector successfully remains on the butterfly. The training scene tasks agents to collectively learn three exercise movements: Horizontal Shoulder Rotation, Forward Arm Raise, and Side Arm Raise.

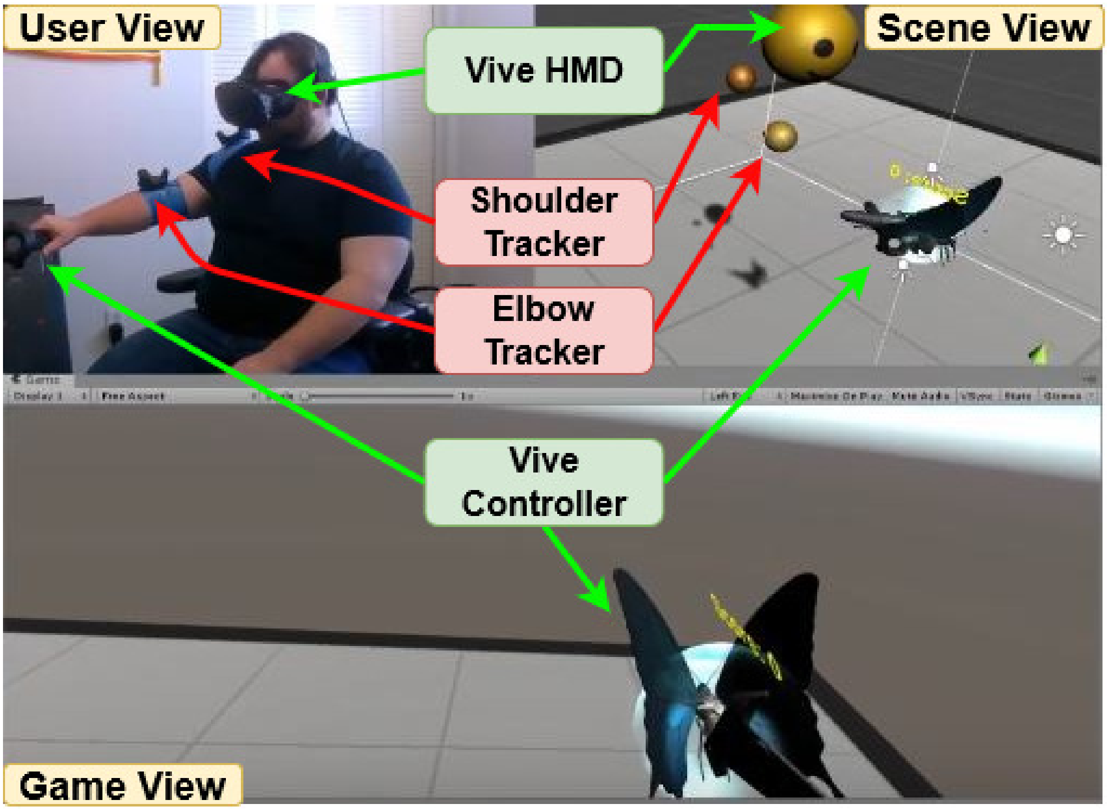


Figure 2.2: Project IB Imitation Learning and User Demonstration. A user demonstrates how to protect a butterfly. Vive Trackers are placed on the user's shoulder and elbow joints to record fixed joint movement dynamics. The agent is set to heuristic control to observe the user's joint torques, angular momentum, and hand (bubble) position. A reward of +0.01 is given to the user per every frame the bubble successfully remains on the butterfly. The recorded demonstration is then used to augment reward during parallel agent training with GAIL & PPO.

Perrin et al demonstrated that bilateral torque enables clinicians to more accurately set guidelines in the rehabilitation of varying athletic groups [313]. Additionally, angular momentum provides a metric to monitor user movement performance over several exercises, ensuring safety and preventing overuse [314]. Several other studies have explored the benefits of quantifying angular momentum for robotic assistance [315], the severity of lower body gait impairment [316, 317], and how it contributes to whole-body muscle movement [318]. Predicting average torque and angular momentum through an AI model may hopefully provide insights for user movements and future assistive robotic design for Project Butterfly to be re-evaluated with exo-skeletal assistance [15, 228].

With our target predictions in mind, we chose to utilize the Unity Ml-Agents Reacher Agent and Deep Deterministic Continuous Control as it observes and predicts agent fixed joint dynamics to complete a given virtual task [307, 308]. We modified the agent to act as a double-jointed virtual arm with specific control and observation on the shoulder, elbow, and end effector joints. This allows our agents to collectively learn from an action space from +/- 1.0 where the agent observes joint torques, angular momentum, and butterfly position to predict shoulder and elbow torque. The agent was given a +0.01 reward per every game engine frame update that the bubble or end effector was successfully on the butterfly. Three exercises were targeted for the agent to learn from Horizontal Shoulder Rotation (HSR), Forward Arm Raise (FAR), and Side Arm Raise (SAR), as shown in Figure 2.3. These movements were chosen as they are considered conventional movement modalities required for active daily living [15, 228].

To examine agent learning, we chose to explore two learning algorithms: Proximal Policy Optimization (PPO) and Generative Adversarial Imitation Learning (GAIL). PPO is a policy gradient method of reinforcement learning that allows sampling paral-

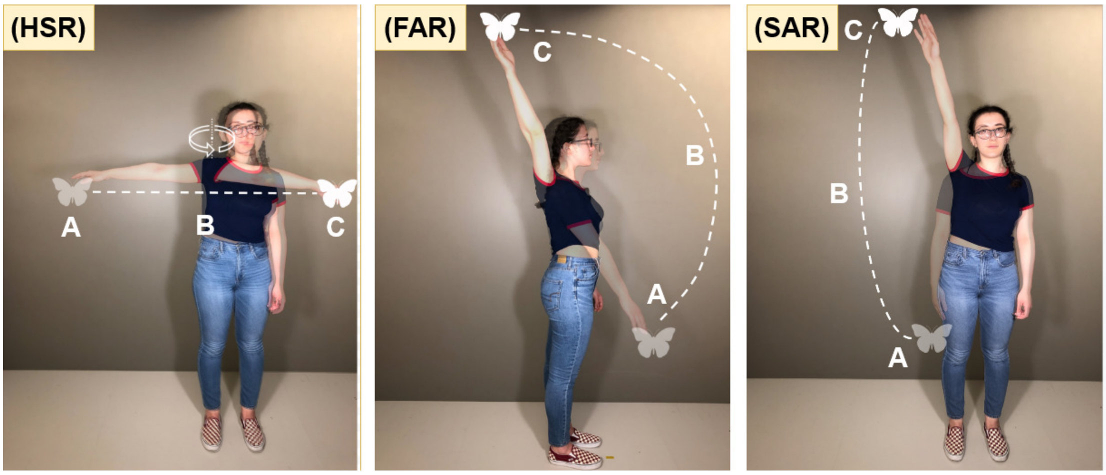


Figure 2.3: Project IB exercise movements for Horizontal Shoulder Rotation (HSR), Forward Arm Raise (FAR), and Side Arm Raise (SAR). Movement directions are indicated by the labels ABC followed by CBA for one repetition.

lel agent interaction with an environment and optimizing the agents objective through stochastic gradient descent [319]. GAIL is an imitation learning method where inverse reinforcement learning is applied to augment the policy reward signal through a recorded expert demonstration [320]. In short, GAIL provides a medium for the agent to imitate the user's exercise, and PPO helps the agent find the maximal reward policy to protect the butterfly.

2.3.2 Agent Training

Two training sessions were examined through Project IB: parallel agent training (as shown in Figure 2.1) with PPO only, and PPO with GAIL. We examined the PPO only model to determine the agent performance when solving for maximal reward and the GAIL + PPO model to see if user demonstrations can influence the training process and or personalize agents to the user's movement biases. For GAIL, a demonstration was recorded for each butterfly exercise movement by a human demonstrator, as shown in Figure 2.2. To record human demonstration, a user was tasked

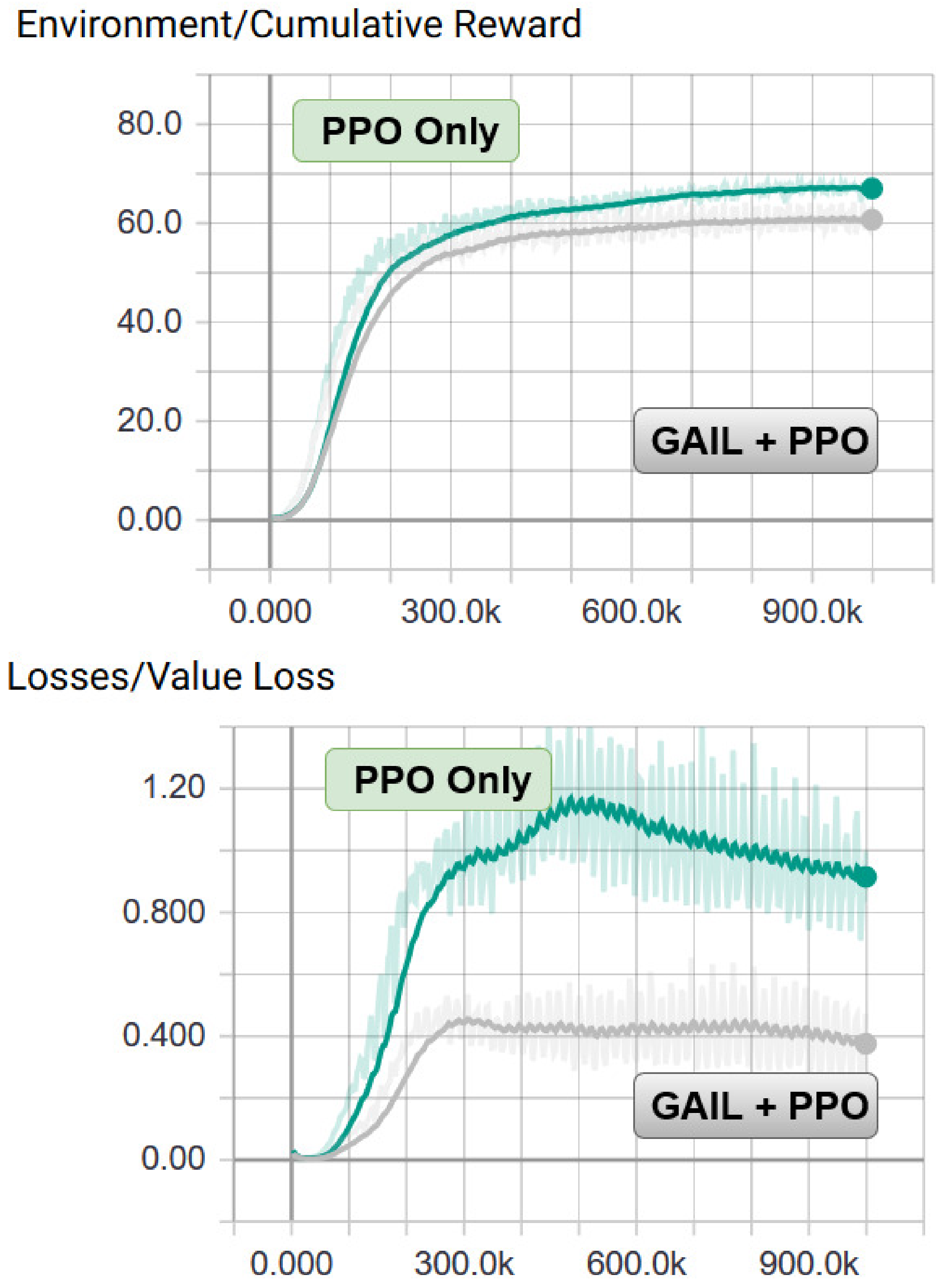


Figure 2.4: Project IB Training Results from Tensorboard for one million steps. Results are viewed from the cumulative 16 agents trained in parallel for the three PBF exercises. The “PPO Only” model attained the highest reward with a 11.4% increase compared the “GAIL + PPO” model. Darker lines indicate smoothed results and lighter lines indicate raw data.

with demonstrating to the agent how to protect the butterfly through arm movement. Vive Trackers were placed at the user's elbow and shoulder joints for agent observation of movement dynamics. This was achieved by creating virtual fixed joints in Unity and inputting rigid body torque and angular momentum into the heuristic agent model. Users demonstrated ideal movements to the agent for about two minutes per exercise.

Training was done with sixteen agents in parallel, as shown in Figure 2.1. Model parameters were tuned to each trainer_config.yaml file as recommended in the Unity ML-Agents v3.X.X plugin [307, 308]. The training parameters differed between "PPO Only" and 'GAIL + PPO," where GAIL was added as a parameter to the PPO reward signal with a strength of 1%. Full tuning parameters and trained models can be found at *https://github.com/avivelor/UnityMachineLearningForProjectButterfly*. Each training model was run for one million steps at a time scale of 100 through the unity ml-agents API. This was equivalent to about a couple hours of training per each model where agents attempted to learn Horizontal Shoulder Rotation, Forward Arm Raise, and Side Arm Raise.

2.3.3 Training Results

Training results between the two models can be seen in Figure 2.4. Both models demonstrated a promising learning rate through one million steps for the 16 parallel agents. However, the "PPO Only" model attained the highest reward with an 11.4% increase compared to the "GAIL + PPO" model. This may imply that the human demonstrator was imperfect in gameplay, and or the motion dynamics recorded through the Vive Tracker require a higher precision. The human demonstrator in Figure 2.2 attained a mean score of 48 between all three movements, which may suggest that the

GAIL + PPO model successfully imitated the user to the best of their ability. While the imitation learning model did receive less reward, the GAIL + PPO model may be useful in understanding user movement bias and weakness. Personalizing agents from user demonstrations may open up pathways to autonomously adjust exercise difficulty around user day-to-day movement capabilities. Subsequently, a future evaluation must be done with a more significant amount of users to understand the ability for personalization and tuning user movement with GAIL as a reward parameter for training.

For the PPO Only model, the deep reinforcement learning alone demonstrated that PPO is highly capable of learning exercise movements by protecting the butterfly. When comparing the results of Figure 2.2 to the Reacher Agent reported by Juliani et al on the Unity ML_Agents Toolkit, the PPO Only model for Project IB received a 41.2% increase in cumulative reward [308]. This may suggest that games like PBF may be an ideal environment for utilizing double-jointed movements, as it was designed for upper-extremity exercise by Elor et al [15]. With the training done, the double-jointed arm for Project IB was then used to provide visual guidance for iVR exercise with PBF. Guidance was done by overlaying the IB Agent as a transparent “ghost arm” as shown in Figure 2.5. With the agents successfully trained, we moved on to perform a small pilot study to see how the PPO Only model competed with human agents.

2.4 User Study

For this study’s scope, we sought to explore how our trained PPO agent would compare to human players. Four users from the University of California Santa Cruz were recruited to compete against the trained “PPO only” model in PBF. Participants were adult college students from UCSC (one female, three males, with a mean age of 23.5

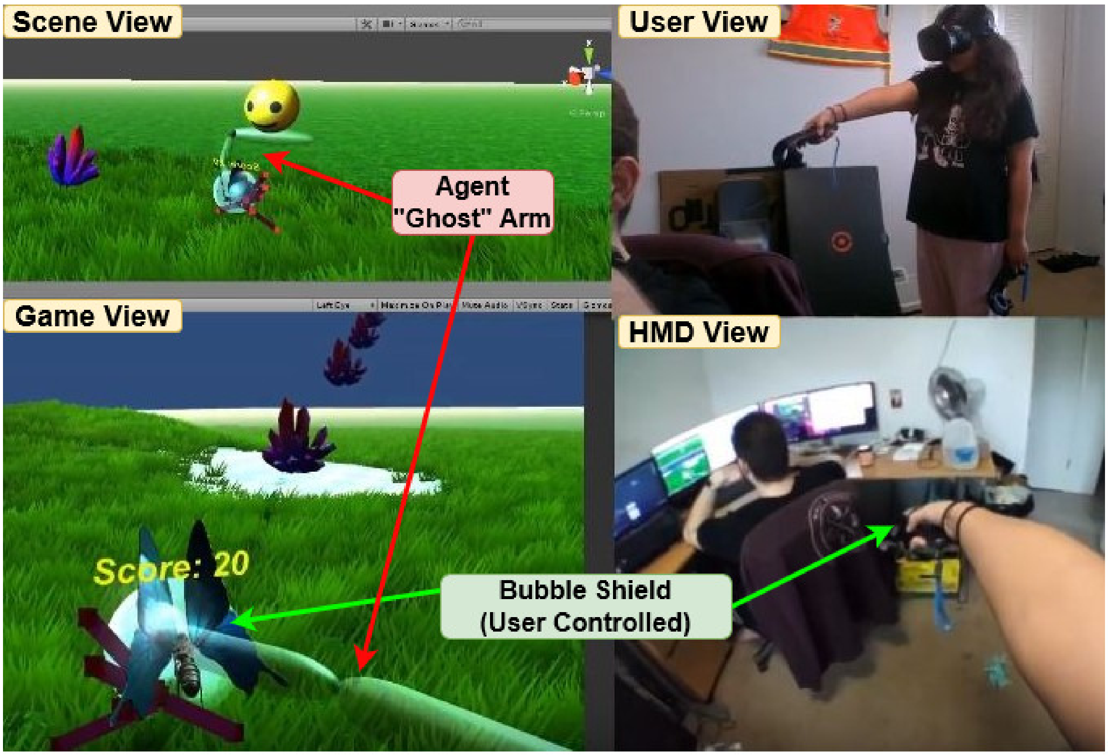


Figure 2.5: Project IB Cooperative Gameplay with Trained Agent. The user controls the bubble shield through the controller as a transparent "ghost" arm appears through the user to help guide and predict user movement in protecting the butterfly.

years old and 1.73 age standard deviation). Each exercise was played for one minute at ten repetitions per minute. A score point is awarded for every crystal the user blocks with the bubble shield on the butterfly. A research administrator was always present to monitor user experience and followed a strict written protocol when interacting with users. Specifically, user testing sessions consisted of the following protocol steps:

1. Preparation: The study administrator sanitized the iVR equipment, made sure all equipment was fully charged, and personally ran a session of Project IB to check the quality of motion capture data communication.

2. Introduction: The administrator instructed the user to remain still and relax. The user was verbally informed about the three exercise movements and the goal of protecting the butterfly. The user was then given a one minute tutorial for each exercise to protect the butterfly with the cooperative IB Agent "ghost arm." An example of this stage can be seen in Figure 2.5.

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

4. Exercise: Users completed 60 seconds of gameplay while competing against the Project IB agent, and the user's final game score was recorded. Upon completion of one set, the Rest stage was repeated. An example of this stage can be seen in Figure 2.6. This stage was repeated until the user successfully completed all three exercises during competition with the agent.

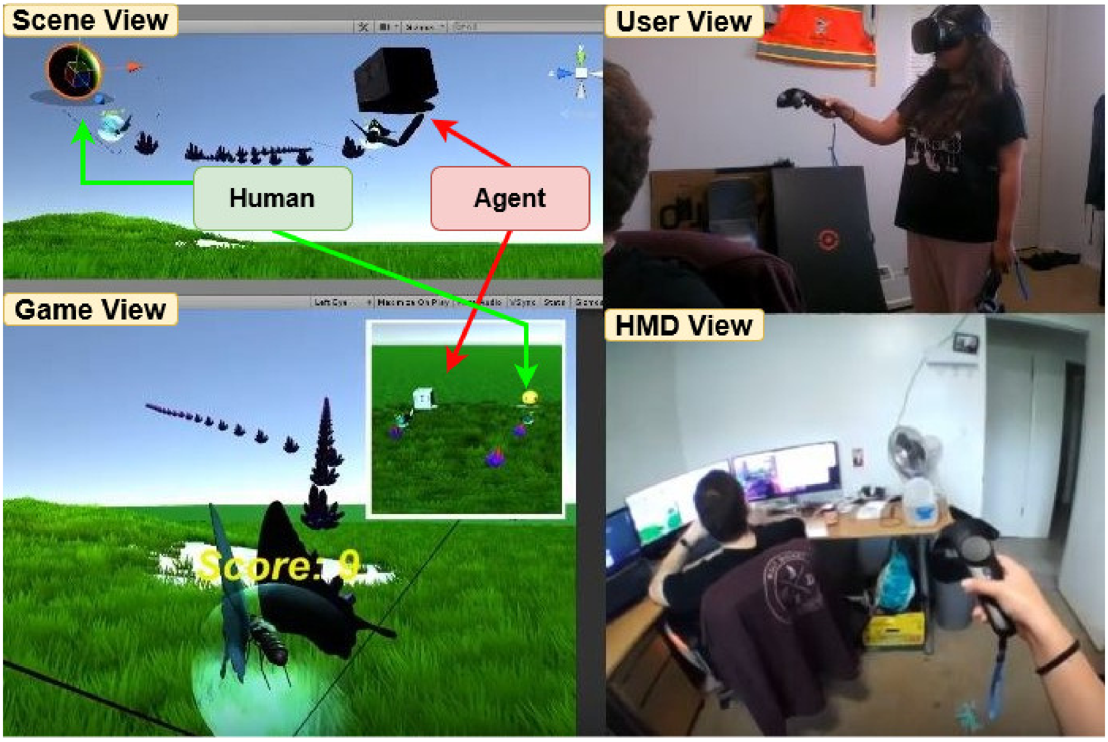


Figure 2.6: Project IB Competitive Gameplay with Trained Agent. The user competes with the Project IB agent to collect the most crystals while protecting the butterfly. The agent is set to the right of the user and is tasked with protecting it's own butterfly. Crystal paths and human vs agent avatar representation are shown in the scene and game view.

2.5 Results and Discussion

Each of the four users from the pilot user study successfully competed with the Project IB agent. The resulting final scores between the users and agent can be seen in Table 2.1. The Project IB agent was able to complete exercises just as well (and even slightly better) than the users for the Horizontal Shoulder Rotation movements. Nevertheless, gameplay indicated that the users were able slightly to outperform the agent for the Forward Arm Raise and Side Arm Raise exercises. Side arm raise appeared to have the highest standard deviation for the agent and the users, indicating a mixed performance. All users reported that they felt the movements were "tiring" at the speed of ten repetitions per minute (requiring a slow and controlled movement in following the butterfly).

While the initial results of Project IB were promising, there are many limitations to consider. More users must compete with both the "PPO Only" and the "PPO + GAIL" models to understand the efficacy of these models as well as exploring unlearned exercises. More demonstrations and imitation learning tuning parameters should be explored with GAIL, such that each model is tailored to each user's movement capabilities for a normalized comparison. Furthermore, a more in-depth investigation must be done to understand the effects of the cooperative "ghost arm" agent to examine if it is assistive from a presence, immersion, embodiment, and self-reported performance perspective. For example, how does the ghost arm compare to the visual guidance from crystals or no guidance at all? These limitations are being considered for future studies with our pilot data in mind.

Exercise	**User Score**	**Agent Score**
Horizontal Shoulder Rotation	46.6 (1.15)	**47.3 (0.58)**
Forward Arm Raise	**45.6 (0.58)**	44.0 (1.00)
Side Arm Raise	**33.3 (4.04)**	31.0 (1.73)

Table 2.1: Results in [mean (standard deviation)] format for human versus agent gameplay. Users were adult college students from UCSC (N=4, F=1, M=3, Age=23.5 +/- 1.73). Each exercise was played for one minute at 10 reps per minute. One score point is awarded per every crystal the user blocks with the bubble shield on the butterfly.

2.6 Conclusion

Through this paper, we presented a novel game mechanic for iVR exercise games that employed deep reinforcement learning and immersive virtual environments to learn from and help guide double-jointed exercise movements. We demonstrated how to convert a previously explored iVR exercise game for machine learning agents. We showcased a methodology of utilizing Generative Adversarial Imitation Learning and Proximal Policy Optimization to exercise with virtual butterflies. We examined two differing models for training our agents, with and without imitation learning. We demonstrated a promising learning rate through training 16 agents in parallel throughout one million steps. We evaluated one of the trained models with a set of four young adults to explore competitive applications with the agent as a game mechanic. The results suggest that with the right training parameters, the model can compete with and adhere to human-level performance in iVR for some exercises after a single training session.

In the future, we hope to explore unlearned exercises and validate a greater range of deep learning models through more extensive user testing to examine its effects on user performance, immersion, and self-reported perception. Our long term goal is to develop an at-home recovery game that uses machine learning to adapt exercise difficulty and assistance. Subsequently, we plan to explore more machine learning algorithms and input parameters such as biofeedback and musculoskeletal simulation to inform of gameplay progression. The incorporation of predictive runtime models to identify muscle weaknesses may further aid in custom movements for an individual user to help maximize their exercise by ensuring the targeted muscles are being used for a given movement. To this end, there are more butterflies to learn from as we continue working towards achieving greater physical intelligence.

2.7 Acknowledgements

We thank Professor Angus Forbes of UC Santa Cruz for his advice during this project and the many participants who volunteered for this study.

Chapter 3

Predictive Shoulder Kinematics of Rehabilitation Exercises through Immersive Virtual Reality

3.1 Summary

Objective: The adoption of telehealth rapidly accelerated due to the global COVID19 pandemic disrupting communities and in-person healthcare practices. While telehealth had initial benefits in enhancing accessibility for remote treatment, physical rehabilitation has been heavily limited due to the loss of hands-on evaluation tools. This paper presents an immersive virtual reality (iVR) pipeline for replicating physical therapy success metrics through applied machine learning of patient observation. Methods: We demonstrate a method of training gradient boosted decision-trees for kinematic estimation to replicate mobility and strength metrics with an off-the-shelf iVR system. During a two-month study, training data was collected while a group of users completed physical rehabilitation exercises in an iVR game. Utilizing this data, we trained on iVR based motion capture data and OpenSim biomechanical simulations. Results: Our final model indicates that upper-extremity kinematics from OpenSim can be accurately predicted using the HTC Vive head-mounted display system with a Mean Absolute Error less than 0.78 degrees for joint angles and less than 2.34 Nm for joint torques. Additionally, these predictions are viable for run-time estimation, with approximately a 0.74 ms rate of prediction during exercise sessions. Conclusion: These findings suggest that iVR paired with machine learning can serve as an effective medium for collecting evidence-based patient success metrics in telehealth. Significance: Our approach can help increase the accessibility of physical rehabilitation with off-the-shelf iVR head-mounted display systems by providing therapists with metrics needed for remote evaluation.

3.2 Introduction

The COVID-19 Global Pandemic has caused an unprecedented need for the advancement of telehealth technologies to provide physical rehabilitation care [321]. 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 [322, 323]. 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 [321]. Yet, for physical therapy to be effectively implemented in telehealth during and beyond the COVID-19 pandemic, current telehealth platforms must obtain evidence-based movement metrics in a remote setting [324].

3.2.1 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 [325]. 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 [326]. Moreover, telehealth has been found to be effective in musculoskeletal practices having demonstrated outcomes and patient satisfaction comparable to in-person care [327]. 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, financial cost, interoperability, the number of connections per session, and additionally built-in features (such as measurement tools, scheduling, playback, libraries, and questionnaires)[328].

3.2.2 Opportunities with Immersive Virtual Reality

Immersive Virtual Reality (iVR) offers an alternative medium to video conferencing with some distinguishing features that meet the criteria established by Cottrell and Russell[328]. Stand-alone head-mounted display systems are becoming more affordable, user friendly, accessible to many users at once, and such virtual experiences can be built with privacy protocols. The systems use low-cost motion tracking methods to match the user movement in the real world to that in the virtual environment [19]. An advantage iVR offers over videoconferencing is the ability for patients and therapists to meet in a three-dimensional virtual environment. With iVR becoming more accessible and providing a more interactive medium to meet patients than videoconferencing, we aimed to investigate iVR in building tools that physical therapists need for remote treatment.

3.2.3 Related Work

During in-person and telehealth sessions, objective assessments that are valid and reliable are a crucial component to diagnose and treat patients [329]. 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 timed get up and go, and balance using the Berg's balance test. For the purpose of this study, 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 include goniometers, motion capture systems, computer vision applications, and sensor fusion techniques using inertial measurement units.

Physical goniometers are the most widely used devices and accurately measure joint angles, but are used for static measurement, and quality is dependent on the tester's level of experience or method for estimating the center of rotation [330, 331, 332, 333, 334, 335, 336]. Virtual goniometers are application or web-based protractors that a user can place on pictures to measure angles and are common tools for telehealth. While virtual goniometers have been validated showing high intrarater and interrater reliability and high agreement between in-person measurements and telehealth measurements [337, 338, 339, 340, 341], they are only able to measure ROM for a single image rather continuously for a video clip. Studies examining knee angle during flexion and walking have included stretch sensor methodologies reporting Mean Absolute Errors

(MAE) ranging from 1.94 degrees to 8.00 degrees [342, 343, 344] and IMU-based estimations reporting errors less than 2.00 degrees [345]. More recently, a review examining inexpensive motion capture systems, predominantly IMU based methods, for measuring joint angles found that shoulder ROM studies reported a MAE ranging from 0.8 to 5.0 degrees [346]. Another inexpensive method uses depth sensors, such as Microsoft's Kinect v2, which has been shown to have measurements as accurate as 3D motion tracking, yet these devices have been discontinued for consumers and are limited to camera based interaction [347]. Pose estimation using computer vision is another technique but has difficulty estimating 3D joint angles with single monocular images [348].

Motion capture systems coupled with biomechanical simulations provide an opportunity to examine this data, as these pipelines are able to determine joint angles and torques using motion capture to inform estimations from musculoskeletal models. There are many software packages on the consumer market to accomplish this (OpenSim, MADYMO, Abaua, LS-DYNA, Aticulated Total Body, AnyBody), which are often used to analyze gait, human motion, injury prevention, and athletic performance. However, the major limitations for using these packages with telehealth is the need for large computing resources to run these simulations as well as complex motion tracking systems that are not often accessible in a home environment. Our goal was to overcome these obstacles by using machine learning paired with off-the-shelf iVR motion capture to predict the inverse kinematics and dynamics that a biomechanical simulation software would generate for the same movements.

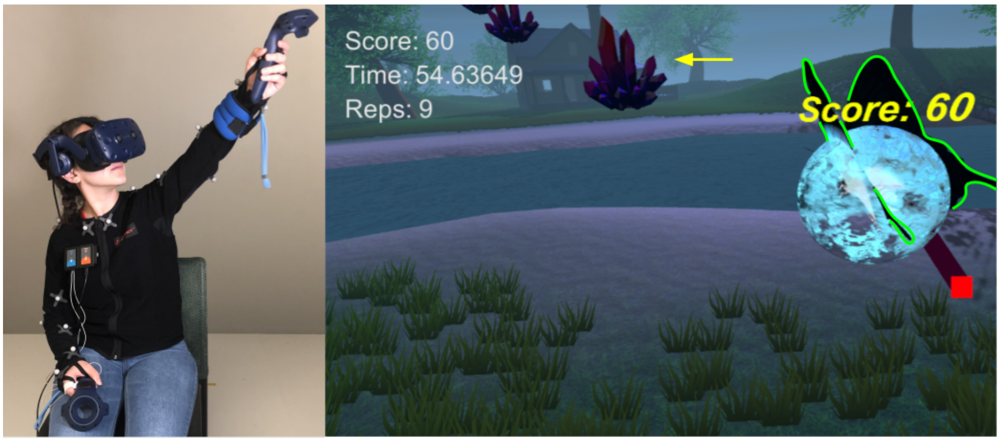


Figure 3.1: A participant is shown playing Project Butterfly using the HTC Vive. The silver dots on the player's upper body are the reflective markers of the motion tracking system, and the blue strap on the arm is a wrist weight to help increase strength. The right-hand image is a capture from gameplay. 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.

3.2.4 Our Contribution

COVID-19 has changed the landscape of telehealth by forcing therapists and patients to adapt to the limited availability of in-person treatment, motivating our lab to develop tools to improve remote rehabilitation. As more insurance companies cover telehealth visits, other disparities are being addressed, including the growing cost of health services, accessibility, and the potential impact of health workforce shortages [349]. An essential step for telehealth and physical therapy rehabilitation is providing therapists with the evidence-based tools needed to remotely evaluate a patient, monitor progress, and effectively communicate. We believe iVR offers an affordable solution that has advantages over traditional videoconferencing solutions. This paper aims to develop and evaluate the feasibility of a machine learning pipeline using solely the motion tracking data of an off-the-shelf iVR system to predict a user's joint angles and torques during exercises within virtual environments.

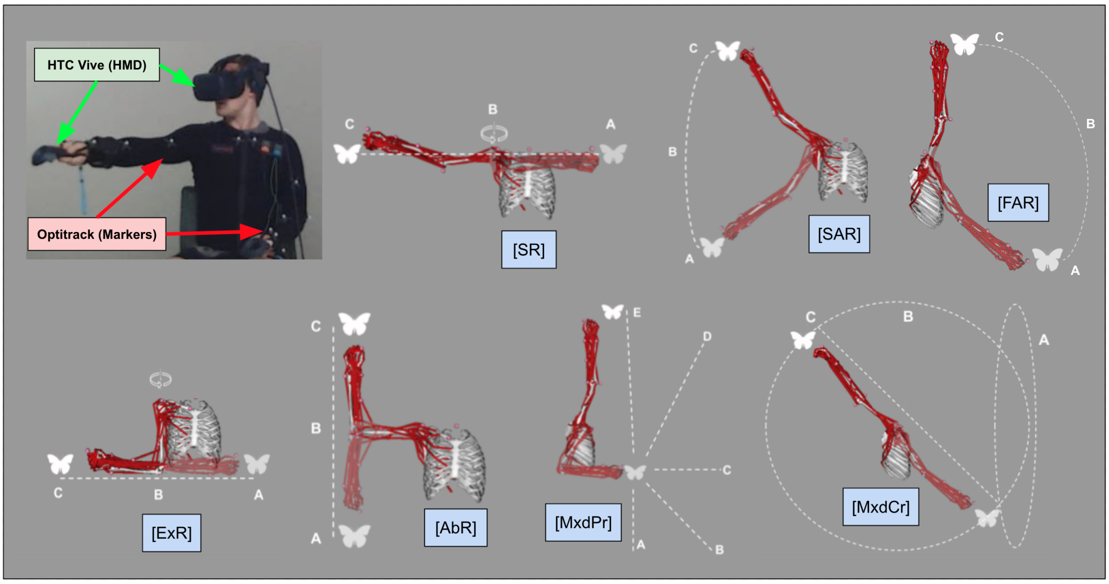


Figure 3.2: Participants played OpenButterfly while seated with ten motion tracking markers placed on bony landmarks as shown in the top left. The game incorporated the seven exercises shown and their abbreviations are in the blue text boxes. The dotted line indicates the flight of the butterfly within 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.

3.3 Methods

The data used to train, validate, and test our model was collected from our previous work entitled “OpenButterfly” [2]. 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 3.1, which received IRB approval from the Office of Research Compliance Administration at the University of California Santa Cruz under Protocol #HS3573. 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 students (one female, four males) with ages ranging from 21 to 28 participated in the study and each student provided informed written consent. Participants continued normal daily living activities during the study. A more complete description of this pilot study can be found in [2].

3.3.1 Protocol

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

- Shoulder Rotation (SR)
- 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. 3.2 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 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 biomechanical simulation 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.

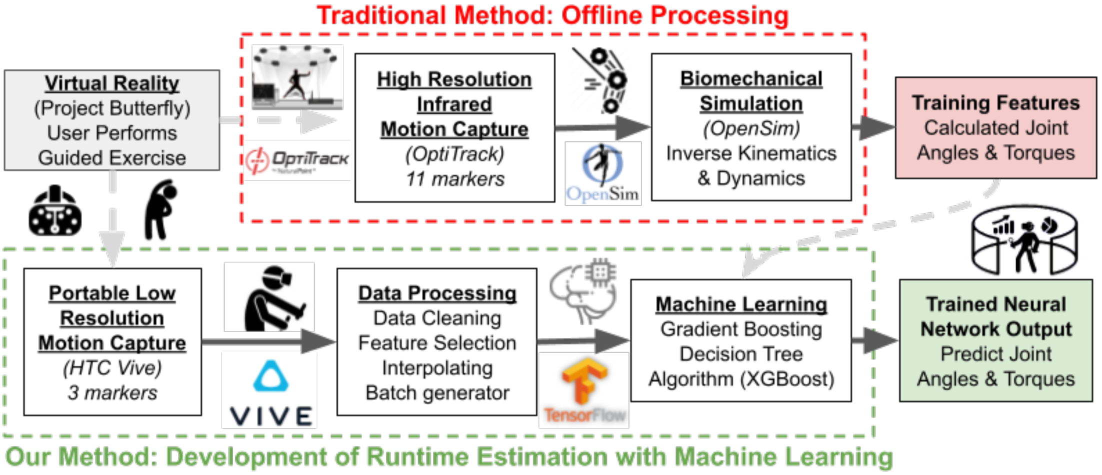


Figure 3.3: Overview of methods to collect data [2], run simulations, and train model. Red pathway shows standard OpenSim method to generate kinematics and dynamics. The green pathway shows our steps to train XGBoost models for predicting the OpenSim results.

3.3.2 Motion Capture and Biomechanical Simulation

Optical motion capture systems are considered the gold standard for accuracy and precision [350], yet these types of systems are expensive and often restricted to laboratory environments [351]. With this consideration, we utilized optical motion capture to collect accurate training data from biomechanical simulation (see the top half of Figure 3.3). 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 [352]. These marker positions are used as input into OpenSim for the Inverse Kinematics Tool, incorporating the upper body model created by Saul et al. [353]. 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:

$$\mathrm{SE} = \sum_{i \in m} w_i \left\| x_i^{\mathrm{exp}} - x_i \right\|^2 + \sum_{j \in \mathrm{uc}} w_j (q_j^{\mathrm{exp}} - q_j)^2 \tag{3.1}$$

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:

$$\mathrm{M(q)\ddot{q}} + \mathrm{C(q, \dot{q})} + \mathrm{G(q)} = \tau \tag{3.2}$$

where

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

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

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

$\mathrm{G(q)} \in \mathbb{R}^N$ 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.

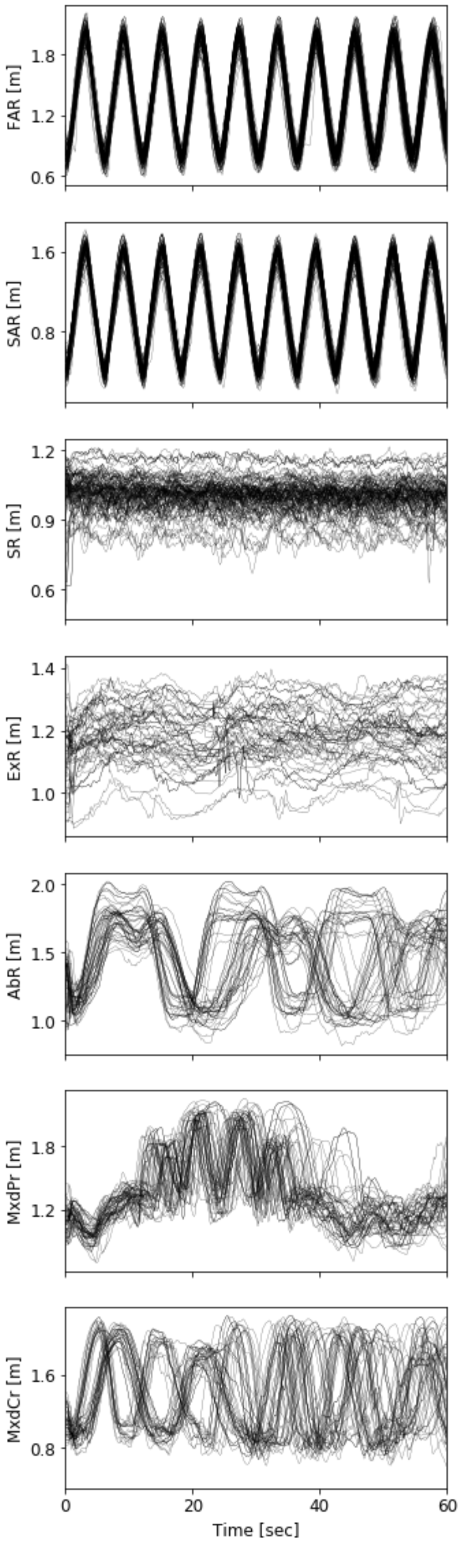


Figure 3.4: Vertical displacement of gameplay controller during each exercise for all users.

3.3.3 Data Analysis

3.3.3.1 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 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.

3.3.3.2 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 [354, 355]. There are pros and cons to each and there is no universal best method for all data sets [356].

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 3.1: Data elements for the machine learning predictive model.

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 [357, 358, 359, 360, 361].

For the purpose of our study, 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 very simple predictive model that has evolved in the machine learning community through many iterative steps including bagging, random forest, boosting, and gradient boosting [362, 363, 364, 365, 366, 367, 368]. 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 [369]. 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 3.1. 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.

3.3.3.3 Model Evaluation

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

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n} \tag{3.3}$$

where

n is number of data points;

y is the prediction of the model;

x is the value obtained from OpenSim.

3.4 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. 3.4 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 3.2. Six models were trained to predict joint angles and torques and Fig. 3.5 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. 3.6, Fig. 3.7, and Fig. 3.8. The MAE comparing the OpenSim results and machine learning models for the unseen test data set is shown in Table 3.3. 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. 3.6 for an entire exercise game of 60 seconds. Additional comparisons are illustrated in Fig. 3.7 and Fig. 3.8 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.

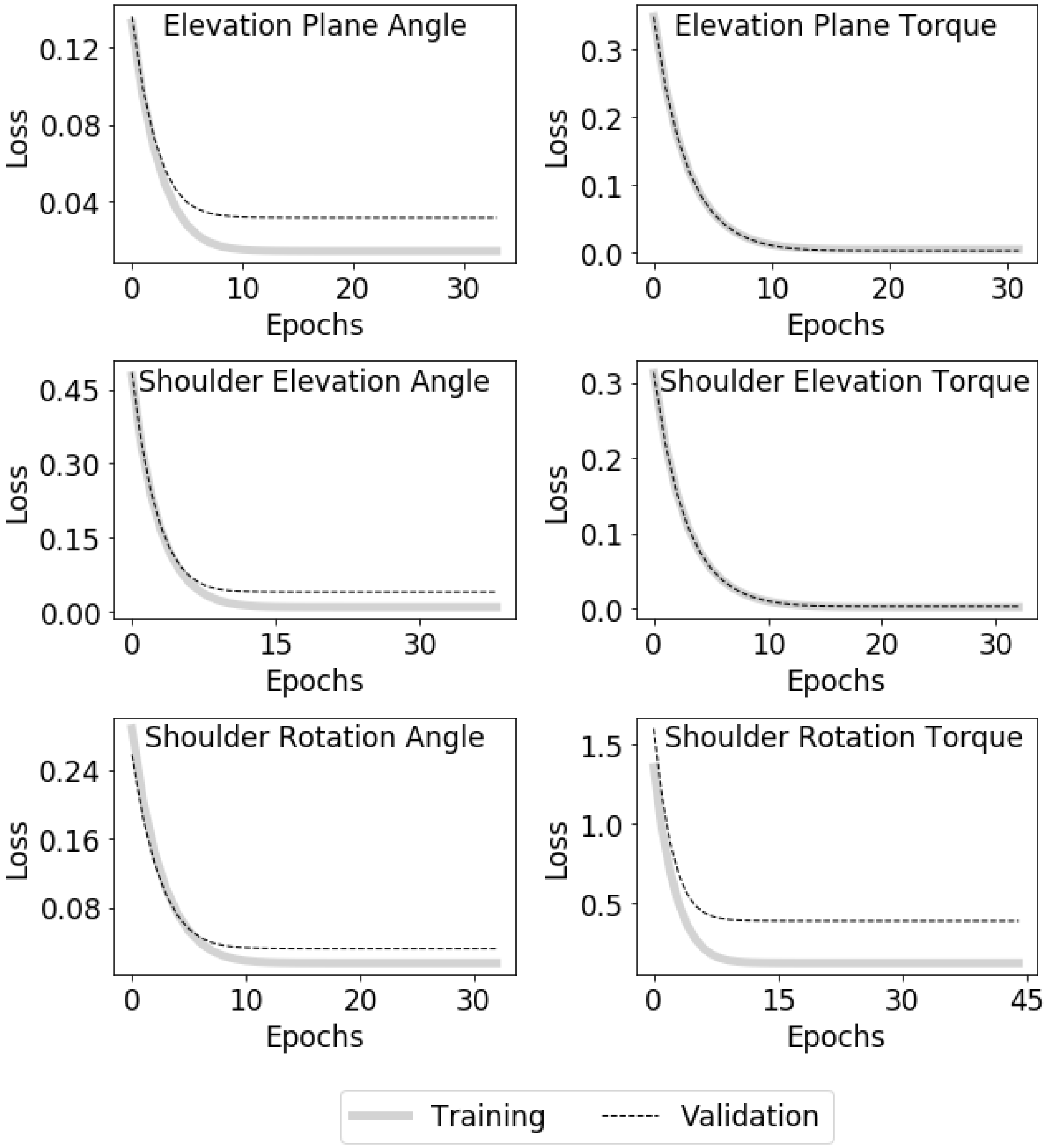


Figure 3.5: Loss function for each model during training to show early stopping preventing over training.

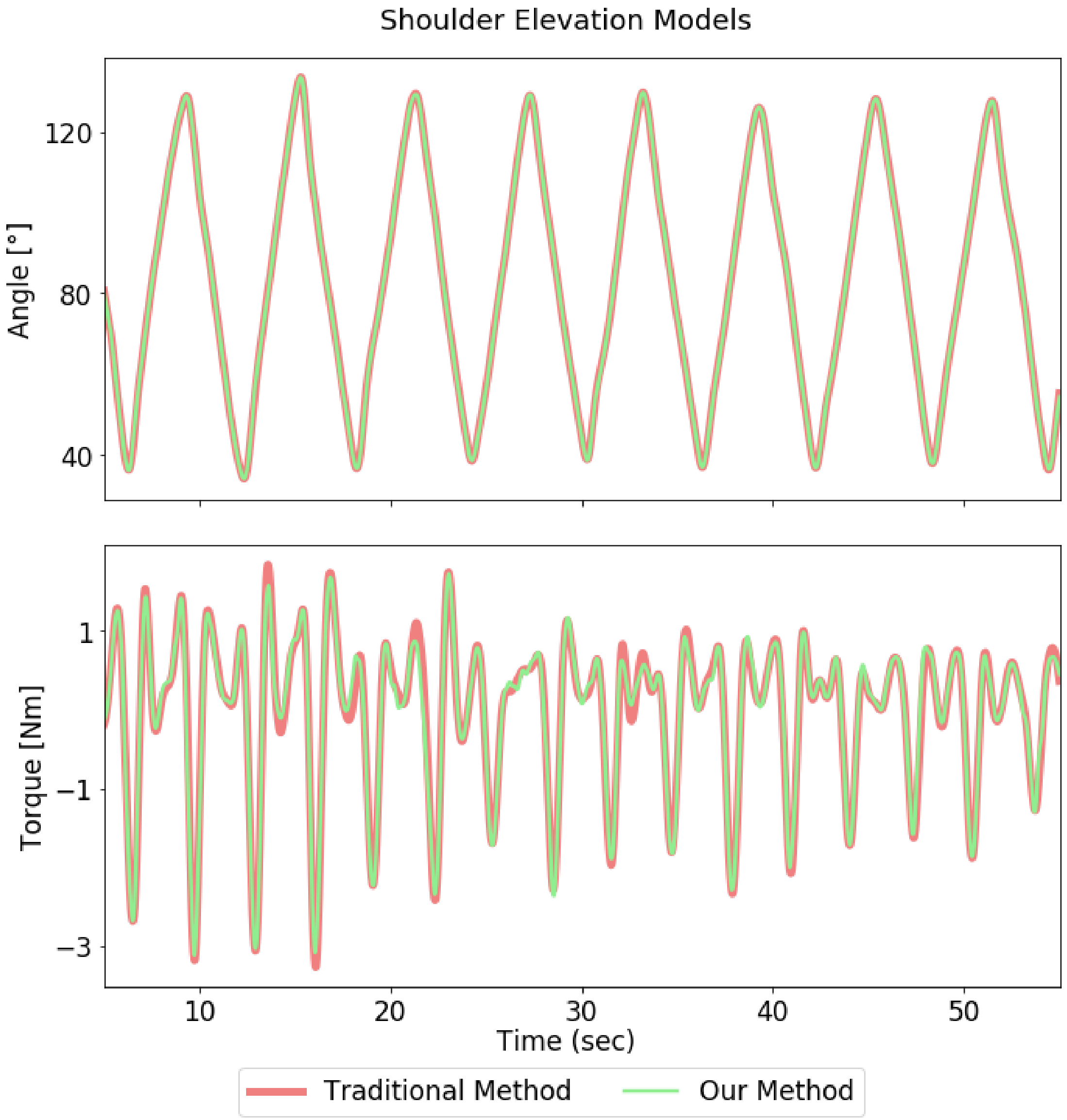


Figure 3.6: An example of OpenSim results and machine learning model predictions for an FAR exercise.

3.5 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

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 3.2: 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 3.3: Mean absolute error between model prediction and OpenSim results for each model's using the unseen test set.

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 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 [370, 371, 372]. 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 [373]. 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 [2]. Additionally, our method can provide dynamic measurements as opposed to static ROM measurements so therapists can monitor smoothness of movement quality as well [347]. These measurements can be provided in run-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 or even used for auto-populating assessment documentation. Run-time 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.

3.5.1 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, 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

Predictive Mobility Metrics of Seven Exercises through ROM [°]

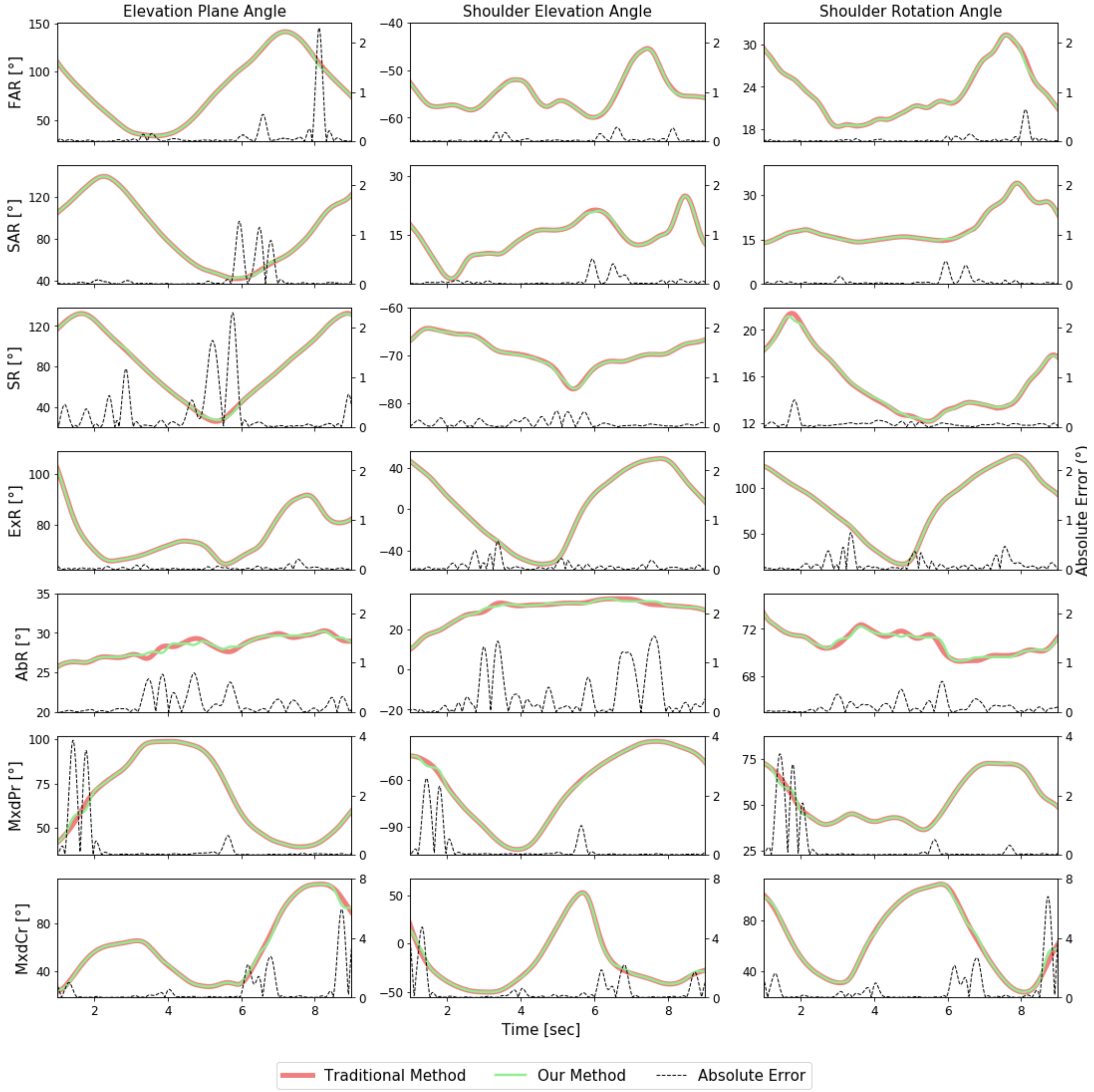


Figure 3.7: Randomly selected segments from the test data set showing the outputs from the traditional method and our method for joint angles for each model with an example for each exercises. Additionally, the absolute error is shown to help see the difference between each method. Exercises are visually demonstrated in Fig. 3.2.

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

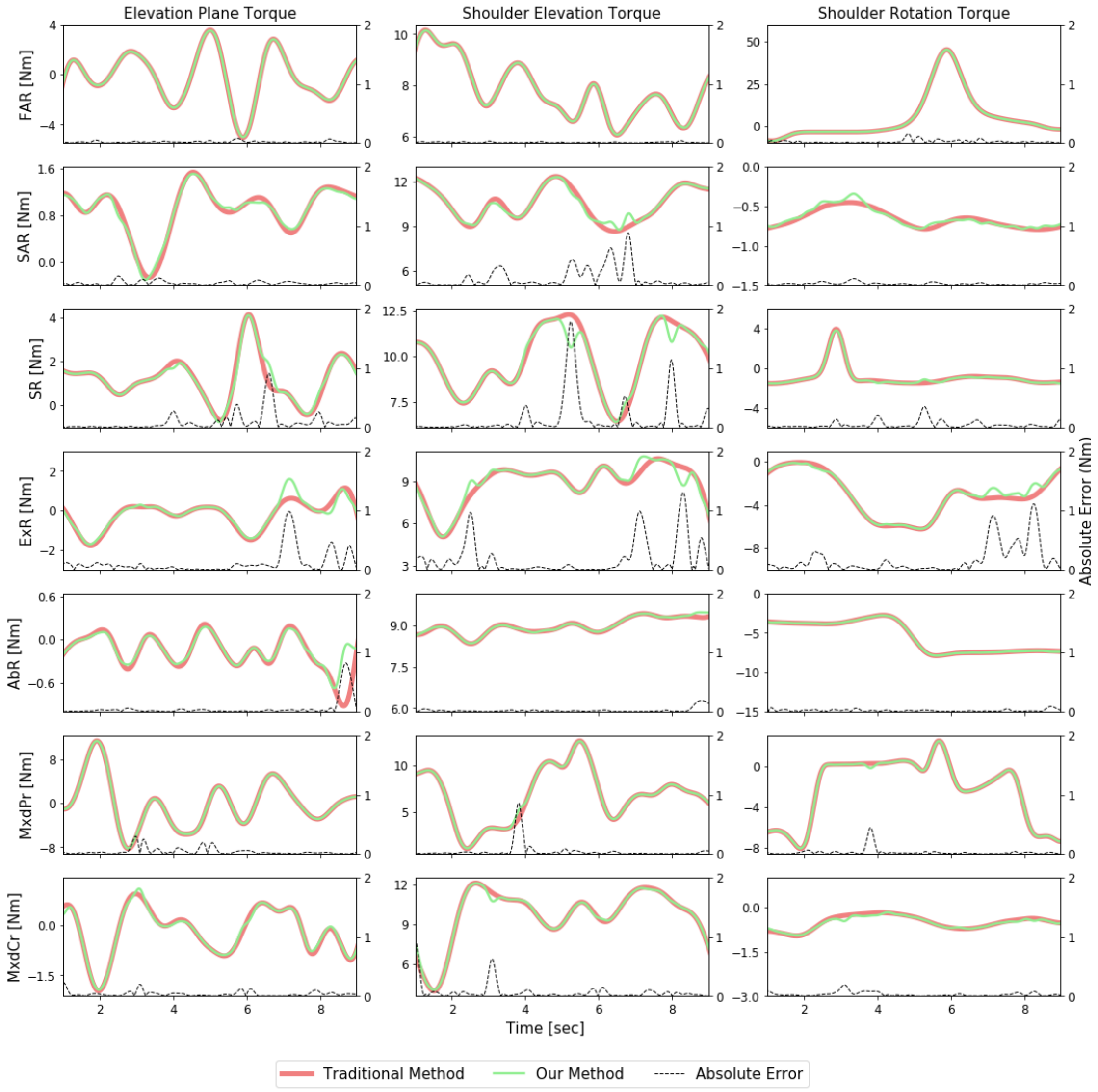


Figure 3.8: Randomly selected segments from the test data set showing the outputs from the traditional method and our method for joint torques for each model with an example for each exercises. Additionally, the absolute error is shown to help see the difference between each method. Exercises are visually demonstrated in Fig. 3.2.

cognitive disabilities, testing soft exo-suits meant for post-stroke rehabilitation, and physical rehabilitation games for users with cognitive disabilities [13, 19, 12, 234]. 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 we would like to 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 [14, 2]. 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 paper.

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.

3.6 Conclusion

This paper 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.

3.7 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 participants of this study, without their time and effort, none of this would be possible.

Part V

ADAPTING IMMERSIVE VIRTUAL REALITY FOR PHYSICAL REHABILITATION

Chapter 1

Introduction

This part examines immersive virtual reality as exercise games and a tool for physical rehabilitation from [227, 22]. Firstly, we explore the efficacy of head-mounted display systems between room scale virtual reality systems for exercise gaming. Secondly, we present a qualitative study highlighting physical rehabilitation clinician impressions and needs for telehealth stemming from 130 interviews at the peak of the COVID-19 pandemic. Third, we investigate the application of immersive virtual reality exercise gaming for shoulder rehabilitation over two months with five participants and follow a variety of therapist recommended success measures from physical, biometric, and game data. This chapter concludes with considerations for future immersive virtual reality exercise games that adopt physical and emotional intelligence design practices in assisting physical rehabilitation.

Chapter 2

On Shooting Stars: Comparing CAVE and HMD Immersive Virtual Reality Exergaming for Adults with Mixed Ability

2.1 Summary

Inactivity and a lack of engagement with exercise is a pressing health problem in the United States and beyond. Immersive Virtual Reality (iVR) is a promising medium to motivate users through engaging virtual environments. Currently, modern iVR lacks a comparative analysis between research and consumer-grade systems for exercise and health. This paper examines two such iVR mediums: the Cave Automated Virtual Environment (CAVE) and the Head-Mounted Display (HMD). Specifically, we compare the room-scale Mechdyne CAVE and HTC Vive Pro HMD with a custom in-house exercise game that was designed such that user experiences were as consistent as

possible between both systems. To ensure that our findings are generalizable for users of varying abilities, we recruited forty participants with and without cognitive disabilities concerning the fact that iVR environments and games can differ in their cognitive challenge between users. Our results show that across all abilities, the HMD excelled in-game performance, biofeedback response, and player engagement. We conclude with considerations in utilizing iVR systems for exergaming with users across cognitive abilities.

2.2 Introduction

Inactivity leads to a decline of health with signification motor degradation: a loss of coordination, movement speed, gait, balance, muscle mass, and cognition [36, 37, 38]. In contrast, exercise is medicine: the medical benefits of regular physical activity prevent motor degradation, stimulates weight management, and leads to a reduction in the risk of heart disease and certain cancers [39]. Nevertheless, compliance in performing regular physical activity is avoided due to high costs, lack of accessibility, and low education [38]. Physical activities are often perceived as chores, especially by those who are not already fit, such as those with disabilities or with musculoskeletal health issues. We posit that the creative use of immersive virtual environments can make physical activities be perceived as fun and enjoyable rather than chores.

Earlier generations of Virtual Reality (VR) systems have proven feasible as effective physical training platforms. For example, it was shown that the Nintendo Wii could be used for physical rehabilitation of older adults with disabilities [374], the Microsoft Kinect was shown to improve balance through exercise [375], and Oculus Rift was shown to be able to help alleviate pain during occupational therapy of pediatric burn patients [294]. However, early head-mounted display systems had significant hardware constraints (e.g., low resolution and low refreshment rates), which leads to non-immersive experiences and motion sickness [51]. Therefore, their use as a health tool was limited. Room sized VR systems, a more costly alternative to head-mounted display systems, were found to be significantly more comfortable and immersive while inducing minimal motion sickness [376]. It was shown that VR systems had some track record of success in promoting exercises for users with rheumatoid arthritis [377] or children with Autistic Spectrum Disorder [378]. All of these studies point to the idea that VR holds potential both as an intervention and as an assessment tool for physical exercise [290]. The power of VR systems lies in the ability of virtual environments to augment user stimuli by conveying various concepts that can be used to produce individualized exercises [290].

Modern immersive Virtual Reality (iVR) systems have gone a long way from a technical standpoint in enhancing user immersion. Such improvements include widening the field of view, increasing frame-rate, leveraging low latency motion capture, and providing a more realistic surround sound experience. Mass adoption of commercial VR HMDs such as the HTC Vive, Oculus Rift, and the PlayStation Morpheus have flooded the market with a combined 200 million systems sold since 2016 [52]. The wide adoption of VR HMDs opens an opportunity to purpose them as a physical exercise platform. This leads to the question of how do we best take advantage of the advanced features that HMDs possess to motivate people from varying walks of life to do physical exercise?

2.2.1 Immersive Virtual Reality and Health

VR holds the unique ability to simulate complex situations that are critical to producing immersive experiences and is auspicious for improving psychologically-based health applications [69]. The use of VR intervention has reported pain-relieving effects when compared to an analgesic during wound treatment [61, 62]. Additionally, VR has been shown to help with Post Traumatic Stress Disorder [63, 64], Borderline Personality Disorder [65], Schizophrenia [68], and various phobias [66, 67, 69].

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 [50]. Thus VR also holds immense potential in physiological rehabilitation as a useful tool for inducing task-based physical exercises [72]. The capabilities of multi-sensory real-time feedback have shown significant outcomes in compliance to exercise protocols [50]. Numerous studies have displayed success in motor improvement from physiotherapy compared to traditional therapeutic intervention [42, 74, 75, 73, 50]. 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 [77, 42]. Thus, there is a need to revisit this examination of VR for health with modern immersive Virtual Reality (iVR) systems [52].

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 [64]. Subsequently, a greater immersion corresponds to a better treatment response, and therefore, is beneficial to improving therapy experiences through virtual environments [91]. Providing engaging stimuli through immersive systems is a crucial factor for the player's experience [70]. The emotional response generated impacts user engagement and helps motivate players to continue with the objectives of the virtual experience [71]. Thus leveraging stimuli to try to instigate a strong emotional response as done in psychotherapy may produce better results in exercise performance. Therefore, a principal research question is: *what is the nature of the relationship between the success of VR stimuli and user emotions?* 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 [133].

2.2.2 Insights from Biofeedback

Biofeedback devices have gained increasing popularity as they use 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 [134, 135]. This activity can be measured through readily available commercial GSR sensors, which has been used to measure the arousal in media such as television, music, and gaming [136, 137]. Cameiro et al analyzed non-immersive VR based physical therapy that uses biofeedback to adapt to stroke patients based on the Yerkes-Dodson law [138], or the optimal relationship between task-based performance and arousal [139]. By combining Heart Rate (HR) with GSR, the authors examined gameplay by quantitatively measuring each user to consider where that 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 [140]. There is a definite potential in evaluating the GSR and HR of each user to determine the intensity of the stimuli between different systems of VR. However, 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 [141]. Modern EEG sensors implemented through Brain-Computer Interfaces (BCIs) have been successful in estimating 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 [142]. 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 [142]. In comparison with 12 other biofeedback experiments, studies that used EEG alone were able to reach 80% maximum emotion recognition [143]. Arguably, the most considerable challenges of BCI are costs, the accuracy of sensors, data transfer errors or inconsistency, and ease of use for devices [142].

Even with these challenges, EEG has been successfully used to understand conditions like Attention-Deficit/Hyperactivity Disorder (ADHD), Anxiety Disorders, Epilepsy, Autism, and Stroke [144, 145]. 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[146]. Through exploring different placements of EEG sensors along with 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 [147]. 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 [141]. From these works, we concluded that there is the potential to analyze brainwaves during iVR stimulus to infer users' emotional responses.

2.2.2.1 On the Subject of Brainwaves

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 [148, 149]. 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 [150]) 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 [150]. **The Beta Band (Focus [151, 152])** 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 [151]. Increases in Beta waves have been correlated to active, busy, or anxious thinking and concentration [152]. **The Delta Band (Awareness [153, 154, 155, 156])** occurs at frequencies between 0.5 to 4 Hertz, and is suggested to relate to awareness and sleep [153]. Delta waves have been found to have the highest activity during deep sleep, where the deeper the sleep, the higher the activity [154]. Researchers have also reported that this frequency band relates to memory interaction [155], such as flashbacks and dreaming [156]. **The Theta Band**

(Sensorimotor Processing [157, 158, 159, 160]) occurs at frequencies between 4 to 7 Hertz, and is associated with sensorimotor processing [157]. This includes spikes in Theta activity for planning motor behavior [158], path spatialization [159], memory, and learning [160]. **The Gamma Band (Cognition [161, 162, 164, 165, 166, 163])** occurs at frequencies between 30 to 100 Hertz, and has been correlated to thought, consciousness, and meditation [161]. Research has theorized Gamma activity is relational to conscious perception [162]. Through studying proponents of meditation and mindfulness training, Gamma activity appeared elevated when a ”conscious experience” would occur, such as shifting mental states in meditation [163]. There are mixed opinions on whether Gamma bands are reliable due to biological artifacts such as eye movement and jaw clenches [164, 165, 166]. However, many researchers argue that Gamma bands show evidence of correlating perception with careful signal processing [163]. 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.

2.2.3 Related Work

Previously, VR HMDs had significant hardware limitations which often caused adverse effects such as motion sickness and eye strain [33]. Thus many immersive media researchers opted to explore alternative VR systems such as room-scale projector-based immersive experiences, which is exemplified in our study by the Cave Automated Virtual Environment (CAVE) [379, 377]. While CAVEs and other room-scale systems have shown great promise, they are an expensive solution – usually in the order of twenty to one-hundred times more costly than modern iVR HMD [377, 52]. Hatada et al suggested that the wider field of view of a CAVE lends to a more immersive experience, with even increased angles as small as 20 degrees vertically and 30 degrees horizontally compared to iVR HMD's inducing a "sensation of reality" [380]. For example, a study exploring acrophobia compared CAVEs to HMD's in 2001 and found that CAVEs created a higher sense of presence and elicited more anxiety [381].

Some past studies have also compared egocentric HMD style devices to CAVE-like systems. Bowman et al did not find significant differences concerning player performance and immersion between HMDs or CAVEs [382]. Philpot et al compared user experience for panoramic video with CAVE and HMDs using interview and survey analysis [383]. The authors found that user responses to thematic relations such as engagement, embodiment, and preference were very similar between the systems as well. Meyerbroker et al used Virtual Reality Exposure Treatment to compare the effectiveness of CAVE and HMD [384]. Similarly to Philpot et al and Bowmen et al, the authors reported no significant differences in effectiveness between the two iVR systems. In a differing example, a comparative study by Kim et al of desktop, CAVE, and HMD system, found the CAVE to be the most performative against HMD-style systems. However, this performance was examined with older HMDs that used a 40-degree field of view [385].

To summarize our literature search, the majority of these past studies either did not find any significance between the CAVE and HMD or found the CAVE to excel. More recent works are suggesting a possible shift in this trend, such as the comparison between Oculus Rift DKII and CAVE2 [386] for collaborative data analysis. Cordeil et al determined that user analysis between the systems did not hold significant differences, yet the HMD based system was suggested to enable faster analysis [386]. The past three years leading up to 2020 have seen a considerable improvement in the HMD technology and consumer adoption, much faster than the CAVE's technology. Pixel density is increasing exponentially in modern devices, increasing more than 30% from 2016 [387]. Given that the issue of affordability can affect public adoption of the iVR system, our study aims to answer the question of whether *modern VR HMDs have finally surpassed room-scale VR systems such as the CAVE in the context of physical exercise.* We aim to answer this unexplored question by examining game behavior, biometric activity, and survey questionnaires.

2.2.4 Study Goals and Impact

This paper reports on a comparative study between the Mechdyne Flex CAVE and the HTC Vive Pro 2018 HMD. We utilize 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. From gameplay, we record each user's game behavior, physical movement, biosignals, and subjective response of gameplay and system use. 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 are as follows:

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.

This paper is among one of the first studies to compare the game behavior, motor movement, and physiological activity of iVR exergaming between room-scale and

HMD systems. We examine these systems with users that have a wide range of cognitive abilities to make sure that our results generalize across a wider range of the population.

2.3 System and Experimental Design

This study uses Project Star Catcher (PSC) [13, 12], 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 [13]. 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 [13].

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 2.1. 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.

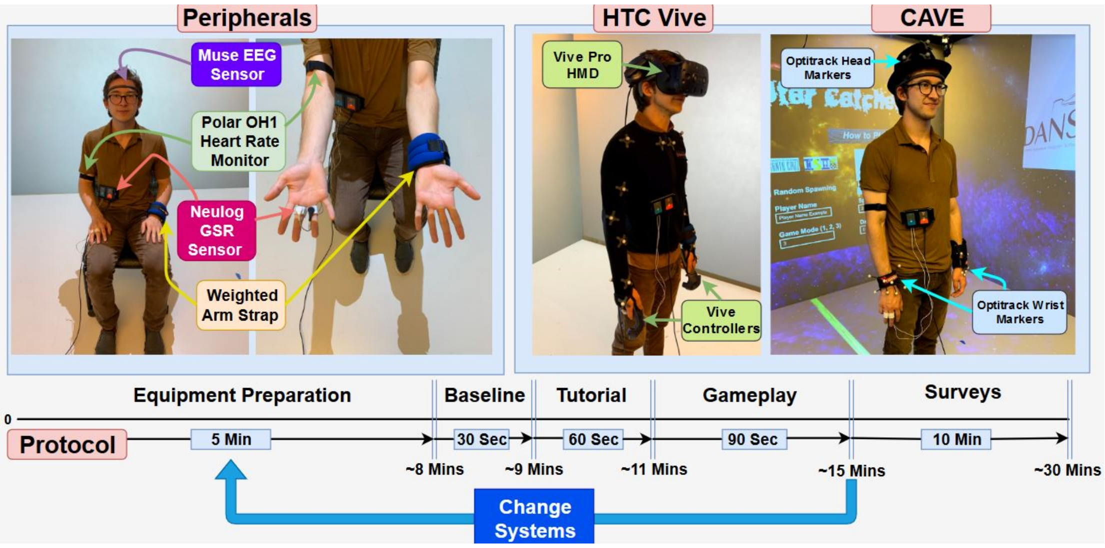


Figure 2.1: System Diagram and Experimental Protocol: Sensor placement (top left), Systems (top right), and experimental protocol (bottom).

2.3.1 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 [388]. 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 can 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.

2.3.2 Data Collection

The following data was collected during the study:

[label=()]HR - Polar OH1 Sensor [389]: an armband with an embedded optical sensor was utilized to wirelessly collect beats per minute by sampling HR activity at 1Hz. GSR - Neulog GSR sensor NUL-217 [390]: 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. EEG - InteraXon Muse 2 - Brain Sensing Headband [170]: collects filtered brainwave data of the prefrontal cortex. The application that communicates with the device uses a Cooley-Tukey FFT to extract waveband power from brain activity [391]. While this headset is relatively low resolution compared to other clinical-grade EEG systems, researchers have used Muse in understanding mindfulness [392], mental states [393], and event potentials [394]. 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 [13]. Motion tracking was achieved with the CAVE through Natural Point Optitrack Motion Capture System [395], while the HTC Vive utilized its native lighthouse localization system for outside-in tracking [396]. Questionnaires - Modified Jennett et al survey [397]: 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 [398, 399].

2.3.3 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 2.1 and is described in detail below:

> [label=()]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. 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. 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. 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. 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. 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 2.1.

2.4 Results

Session data was post-processed using the Mathworks MATLAB 2018b Statistics and Machine Learning Toolbox [400]. 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 [401]. 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 [382, 384, 385, 386, 383]. We discuss these findings in the following subsections.

2.4.1 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 2.3. 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 2.4 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 2.5 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.

2.4.2 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 this paper, 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 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 2.1, and indicate no significant difference between pre-gameplay states between systems, with the exception of the gamma band for the group without disabilities.

Figure 2.6 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 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 2.1 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 2.7. 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 2.6.

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

	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 2.1: 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.

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 biometric performance of our users.

2.4.3 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 [397]. 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 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 2.8, 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 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 2.3 and 2.4). 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.

Self-reported emotions felt during gameplay can be seen in Figure 2.9. 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 2.9. 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 2.7. 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.

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 2.10 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 2.10, 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].”

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.”

2.5 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.

2.5.1 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:

2.5.1.1 *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 [13]. Another one of our previous studies has explored iVR games to strictly follow exercises by protecting a "cute" virtual butterfly in "Project Butterfly" (PBF) [234]. 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.

2.5.1.2 *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.

2.5.1.3 *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.

2.5.2 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 [61, 62]. 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 [13]. 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 [386]. However, we argue that modern commercial HMD based systems have fully surpassed the advantages the CAVE used to hold with 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 off the shelf systems like the

HTC Vive to show full-body motion capture approximation and ease of implementation [402]. Furthermore, there is a vested interest in producing full-body motion capture by industry competitors for the future of iVR interaction with HMDs [403, 404]. 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 [52]. 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.

2.5.3 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 [13, 234]. 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 [405, 392, 393, 394].

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 paper 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.

2.6 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 [29]. 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" [32]. Modern iVR systems are enabling deeper and more rich experiences of presence into the virtual world [52]. These elements of the ultimate display that influenced modern iVR as we know it appears to be near at hand [406]. Given this progress of immersive mediums into the virtual world, we ask: *what would the ultimate iVR system be for exergaming and health*?

This paper 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 [13]. 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 runtime emotional and physical state of each user to create a profound and maximally engaging experience. Subsequently, there are far more stars to catch and biofeedback to record as we collectively build our vision of the ultimate immersive display for healthcare and exergaming.

2.7 Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant #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 Hope Services California and the many caretakers, health-care professionals, and student participants that helped make this study possible. Additionally, the following UC Santa Cruz students are acknowledged for their support: Ruchi Gupta, Tiffany-Ellen Vo, Pavel Frolikov, and Conrad Esch.

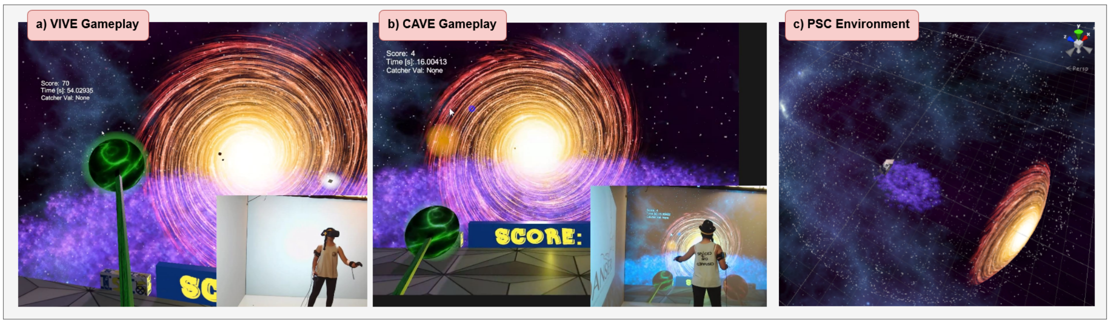


Figure 2.2: System Gameplay: a) A user catches a shooting star with the HTC Vive. b) A user prepares to catch a shooting star with the CAVE. c) The PSC virtual environment.

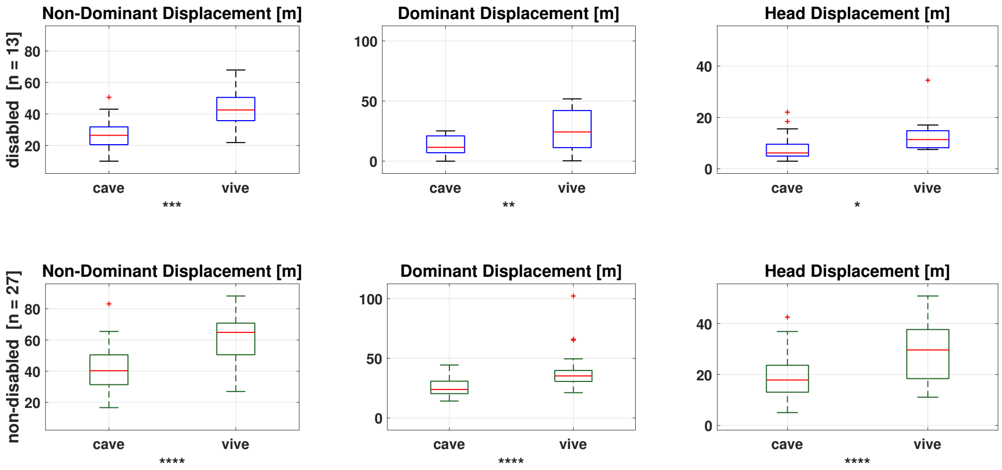


Figure 2.3: Player movement of users with disabilities (row one) and without disabilities (row two). Wilcoxon significance level between CAVE and VIVE is indicated in asterisk notation and "ns" indicates not significant (highlighted in red). Note that Non-Dominant Displacement indicates the total movement of the weighted arm during Project Star Catcher Gameplay between systems.

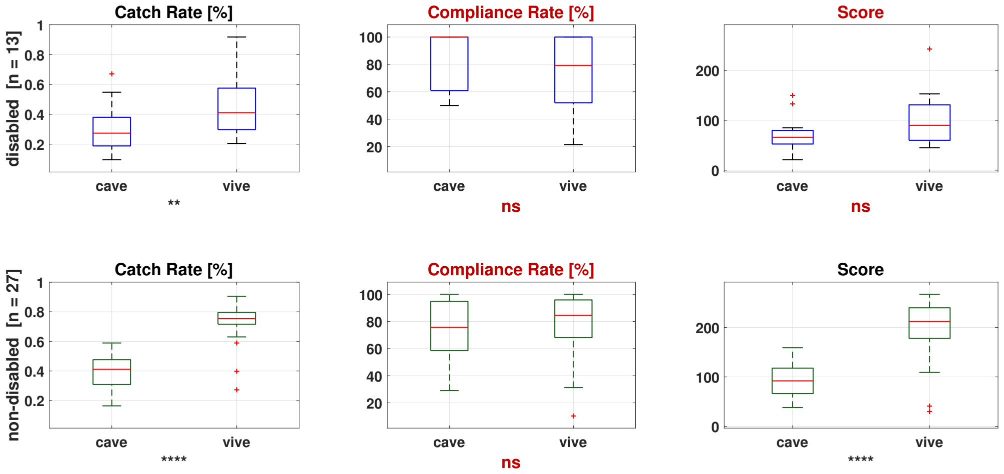


Figure 2.4: Gameplay score and success rates of users with disabilities (row one) and without disabilities (row two). Wilcoxon significance level between CAVE and VIVE is indicated in asterisk notation and "ns" indicates not significant (highlighted in red).

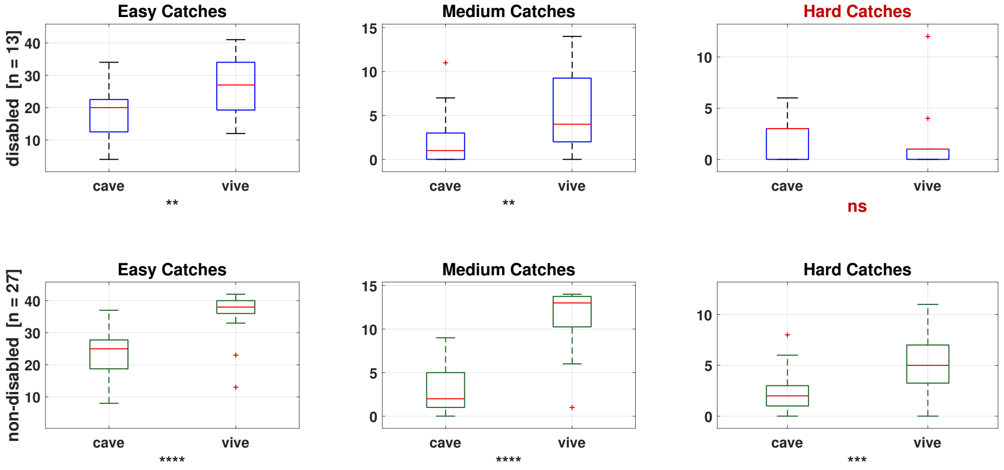


Figure 2.5: Successful star catches with difficulty of users with disabilities (row one) and without disabilities (row two). Wilcoxon significance level between CAVE and VIVE is indicated in asterisk notation and "ns" indicates not significant (highlighted in red).

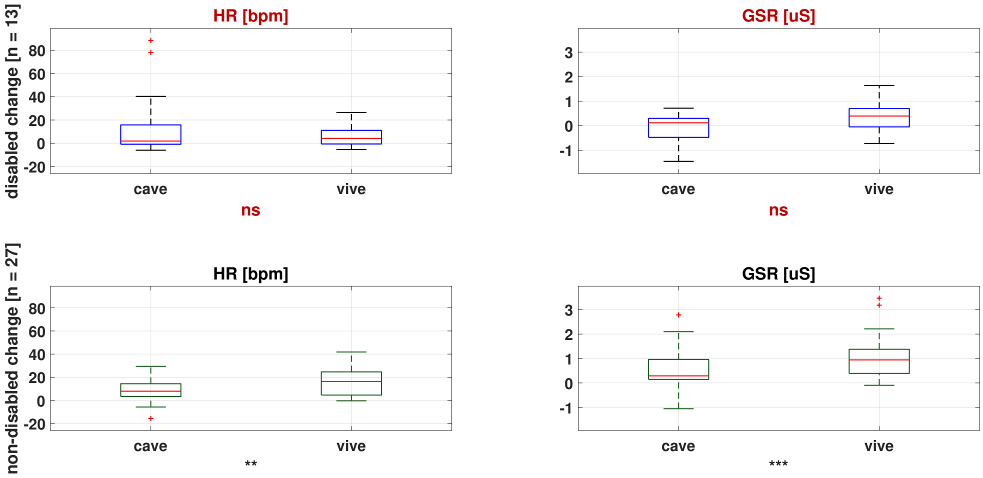


Figure 2.6: HR (in beats per minute) and GSR (in micro-Siemens) resting state change from gameplay of users with disabilities (row one) and without disabilities (row two). Wilcoxon significance level between CAVE and VIVE is indicated in asterisk notation and “ns” indicates not significant (highlighted in red).

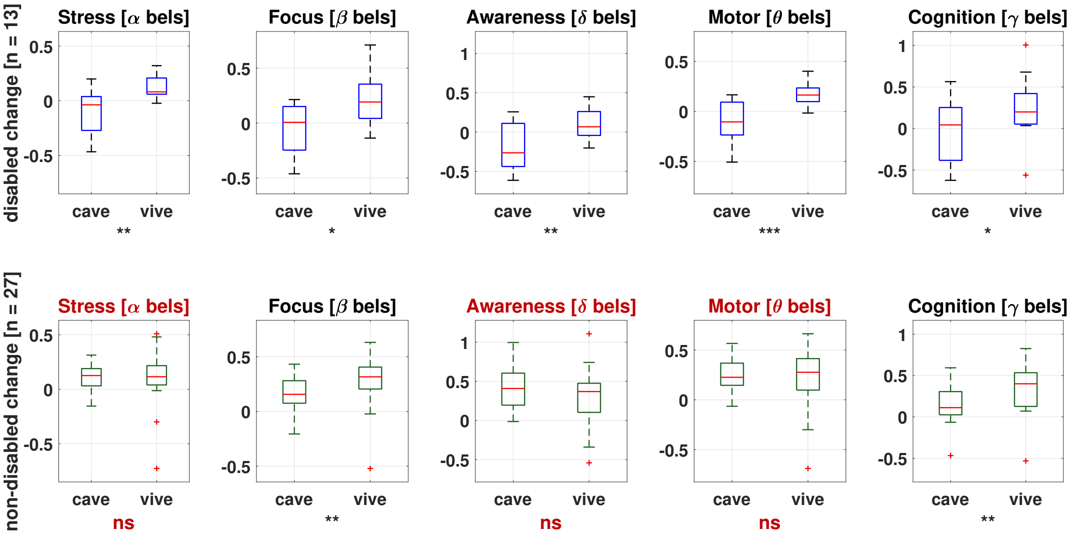


Figure 2.7: EEG brainwave power in bels from resting state change induced during gameplay for users with disabilities (row one) and without disabilities (row two). Note that stress, focus, awareness, motor, and cognition are represented by the alpha (α), beta (β), delta (δ), theta (θ), and gamma (γ) band powers. Wilcoxon significance level between CAVE and VIVE is indicated in asterisk notation and “ns” indicates not significant (highlighted in red).

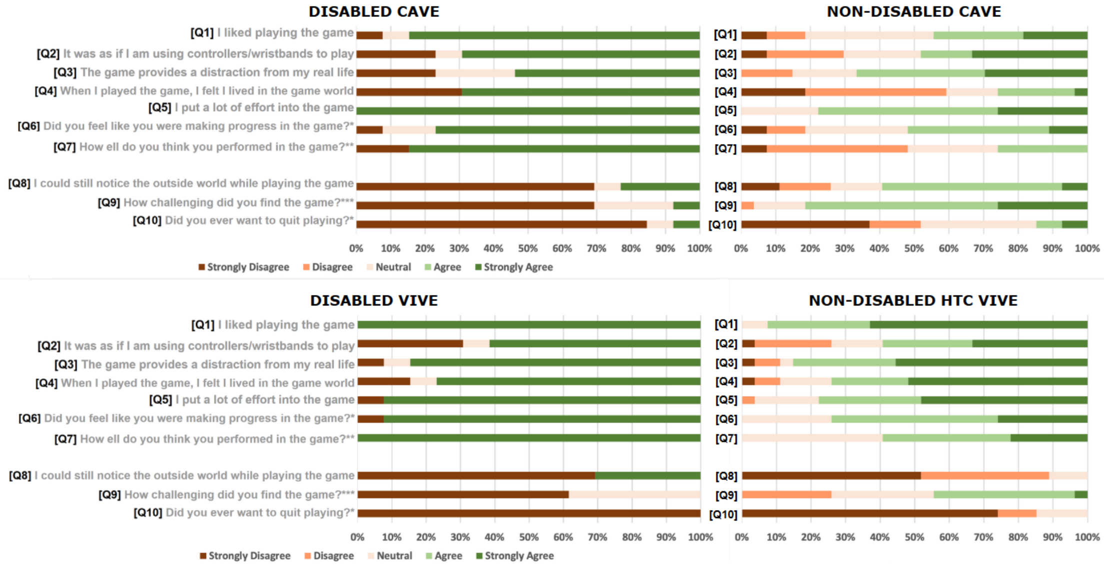


Figure 2.8: Subjective rating questionnaire responses for the between user groups and systems. For Q1-7, strongly agree is the desired outcome. For Q8-10, Strongly disagree is the desired outcome. Disabled user responses were modified to 3 point scale as recommended by healthcare professionals from Hope Services, CA, to increase accuracy. * = "Not at all" to "a lot," ** = "Very poor" to "very well," *** = "Not at all" to "very challenging."

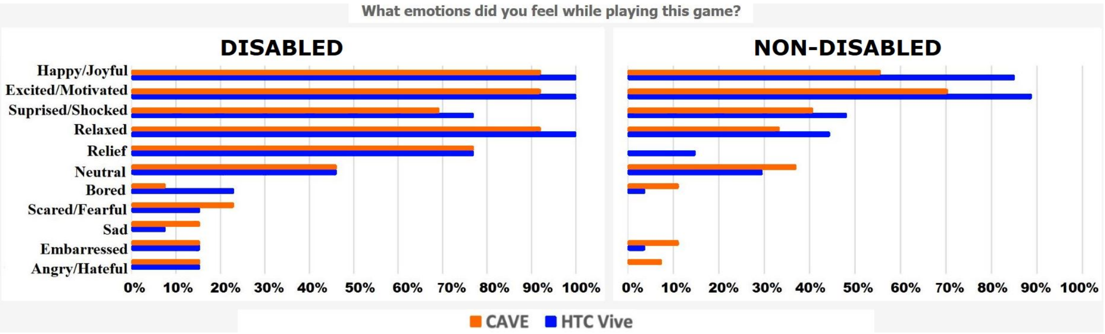


Figure 2.9: Self-reported emotions strongly felt between the two different systems and user groups.

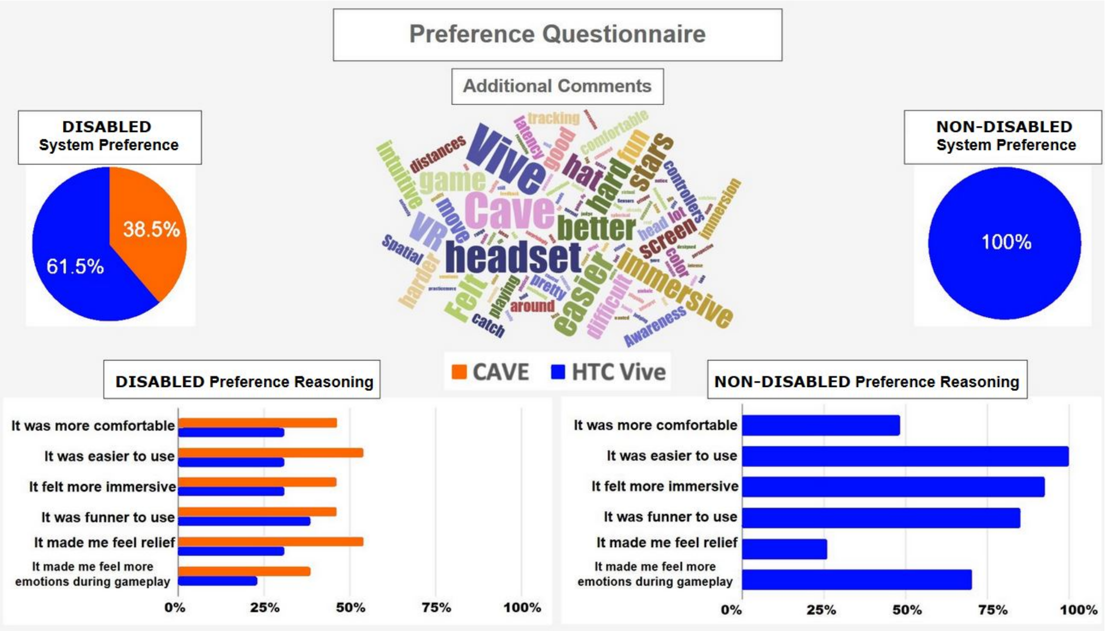


Figure 2.10: System preference between the two user groups with reasoning for preference.

Chapter 3

A Newfound Virtual Reality in Physical Rehabilitation: A Qualitative Study on Therapist Impressions of Telehealth and Technology Needs Amidst the COVID19 Pandemic

3.1 Summary

IMPORTANCE: Most physical therapists would agree that physical rehabilitation is difficult to perform remotely. However, COVID-19 has forced many physical therapists and their clients to use video conference telehealth. With widespread use, we learn the shortcomings of current methods and how to steer future technologies to promote remote patient evaluation and rehabilitation. A key conclusion is that many therapists believe that telehealth will continue to be an integral component of their practice. OBJECTIVE: We interview physical therapists to understand the strengths and weaknesses of their current solutions and what they believe would help the patient rehabilitation process with telehealth. DESIGN, SETTING, AND PARTICIPANTS: We interviewed 130 physical rehabilitation professionals across the United States through video conferencing during the COVID19 pandemic from July - August 2020. Interviews lasted 30-45 minutes using a semi-structured template developed from an initial pilot of 20 interviews to examine potential barriers, facilitators, and technological needs. MAIN OUTCOMES AND MEASURES: Physical therapists have lost their ability to feel their patients injuries, easily assess range of motion and strength, and freely move about to examine their movements when using telehealth. This makes it difficult to fully evaluate a patient and many feel that they are more of a "life coach" giving advice to a patient rather than a traditional in-person rehabilitation session. RESULTS: The most common solutions that emerged during the interviews include: (1) immersive 3D technologies which allow PTs and clients to remotely walk around each other, (2) evidence-based measures, (3) automating documentation and measurements, and (4) clinical practice operation through the cloud. CONCLUSION AND RELEVANCE: Many physical therapists believe that telehealth will continue to be an important form of rehabilitation, despite difficulties with current methods to provide quality care.

3.2 Introduction

n 2020, the global COVID-19 pandemic halted existing models of physical rehabilitation care in the United States [407]. Rehabilitation clinics were disrupted with social distancing guidelines, new protocols for patient care, and reformed technical infrastructure for remote care [408]. As a result, the global pandemic has deepened existing inequities and gaps in modern healthcare [409]. These challenges sparked widespread adoption of telehealth, a service once limited for physical rehabilitation due to its lack of hands on evaluation capabilities for physical therapy. The American Physical Therapy Association (APTA) notes “Telehealth is a well-defined and established method of health services delivery... [and supports] Advancement of physical therapy telehealth practice, education, and research to enhance the quality and accessibility of physical therapist services [410].” Subsequently, the practice of telehealth for physical rehabilitation has seen an explosive adoption to mitigate the impact of the pandemic. In this paper, we examine the effects of telehealth adoption: it’s challenges for physical rehabilitation, successes and practices during the pandemic, and professional views of how telehealth will transform physical rehabilitation beyond the pandemic.

3.2.1 Telehealth and Physical Rehabilitation

In 2000, VandenBos et.al. defined telehealth as a “real-time service that occurs when the patient and the provider are physically separated at the time the service is rendered and some communication device is used in the exchange between the parties [411].” The best communication device then was a telephone. With advancements and accessibility of video camera technology, the internet is a higher choice for telehealth visits.

In recent years, telehealth has proven to be an effective way to provide certain medical services. In a study by Axelsson et al, cognitive behavior therapy delivered via the internet “appeared to be noninferior to face-to-face [cognitive behavior therapy] for health anxiety, while incurring lower net societal costs.” Moreover, online therapy has the capability to increase evidence-based treatment [412]. Another study in 2016 established that telehealth has potential to provide more accessible services to “vulnerable patient groups,” especially those who have transportation or financial barriers [413].

Similarly, Telehealth Physical Therapy (TelePT) has been making its way onto online services; however, it is uncertain if telehealth can be truly effective for physical rehabilitation. First, what is TelePT? Building upon VanderBos, TelePT is a service that occurs when the patient and the physical therapist are physically separated at the time the service is rendered and some communication device is used in the exchange of rehabilitative exercises and consultation. The growing effectiveness of TelePT is supported by ongoing research. A study from Northeastern University examining virtual environment based systems for upper extremity mobility rehabilitation for post stroke patients found improvements even after physical therapy was discontinued [414]. Another study recorded upper and lower extremity movements with images and sent them to a hospital whose physical therapists provided feedback. This study determined that tech for rehabilitation was effective for storing data and referencing back later.8 Reed et. al. compared TelePT at home versus in-clinic rehabilitation and found that “telerehabilitation has the potential to substantially increase access to rehabilitation therapy on a large scale [413].” These studies show the results for online physical therapy have been effective with long term benefits to both the patient and the therapist.

3.2.2 Telehealth in the COVID19 Era

While telehealth was not initially the norm, it’s use skyrocketed when people were required to abstain from physical contact with non-household members. During the 2020 pandemic, Smith et. al. argued that telehealth requires “a significant change in management effort and the redesign of existing care models of care. Implementing telehealth proactively rather than reactively is more likely to generate greater benefits in the long-term, and help with the everyday (and emergency) challenges in healthcare [415]." The pandemic forced many providers and small health businesses to switch to virtual visits as essential elements of access and wellbeing.

Dantas et. al says that TelePT "offers the possibility to continue providing some physical therapy services to patients, but regulations and implementation barriers are extremely heterogeneous around the world [416]." Alan C. Lee also notes the benefits of using TelePT for building strong relationships between the patient and provider while maintaining a safe environment [417].

Telehealth cannot replace all forms of medical care. COVID-19 has deepened struggles for patients who require constant treatment or cancer patients who have compromised immune systems [418]. While it is a significant resource with growing adaptations, Telehealth can not provide blood withdrawal, chemotherapy, or physical examinations.

Still, COVID-19 has elevated the relevance of TelePT and its efficacy. As technology remains a huge part of daily life, TelePT might offer new opportunities for regular care practice. Remaining challenges to facilitate adoption include the need for standardized telehealth practices and studies that may address the efficacy of telehealth compared to in person which address what is lost without manual therapy. Furthermore, the uncertainty of insurance coverage for TelePT may be a significant barrier.

3.2.3 Research Goals

In this study we find rehabilitation via telehealth offers sustainable and digitally accessible community based resources especially for underserved communities.

While research on telehealth telepresence for physical wellbeing are largely reviews or perspectives, this article is based on primary research across the whole country. Furthermore, we have sought direct feedback from physical therapists (PT), a group that is often not heard in telehealth research. While physicians may make diagnoses using video, PTs are impacted the most as they lose hands-on evaluation of injuries for rehabilitation. COVID and social distancing protocols have limited their access. Video therapists and TelePTs continue to practice in this global pandemic. It is critical to learn what works, or not, in moving forward beyond the global COVID pandemic. The niche transformation from therapy to rehabilitation medicine is crucial to adoption. This key finding reflects the significance of our research towards making an intervention in existing literature.

3.3 Methods

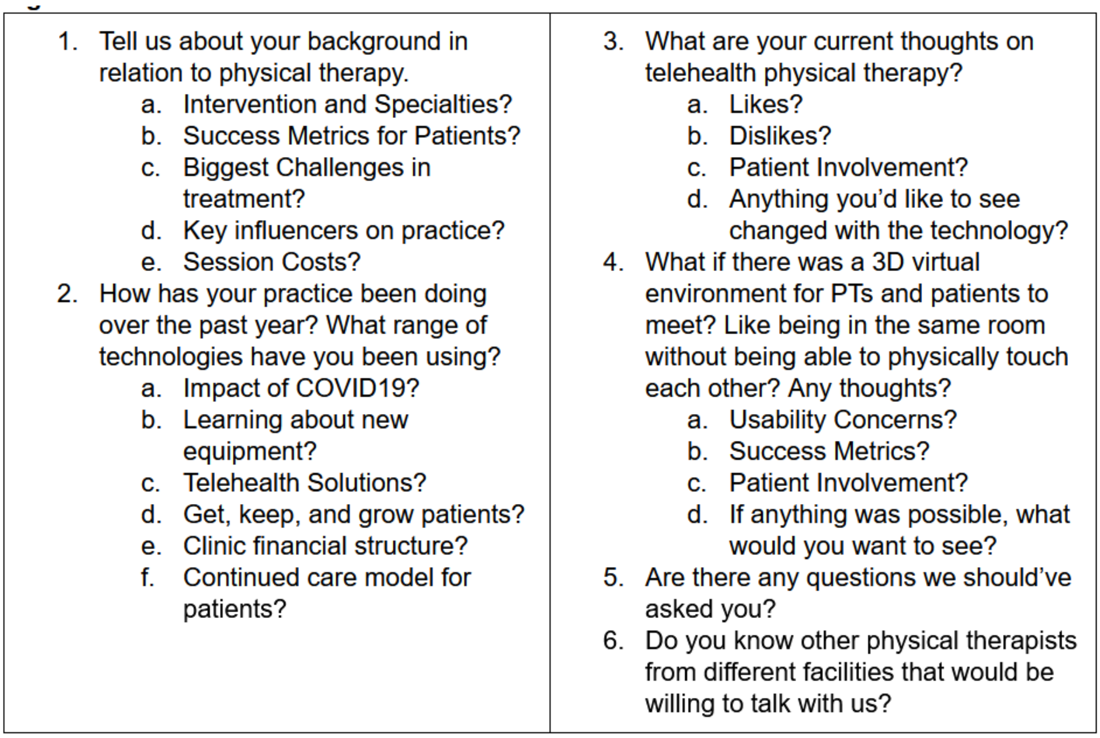
3.3.1 Study Design and Sample

This study used a 2-stage sampling design. First, we purposefully selected 3 states (California, New York, and Florida) with the largest spikes of COVID19 cases to start. Within each state, we identified regional chapters of the American Physical Therapy Association (APTA) and reached out to professionals using LinkedIn, an online business social network platform. Two researchers identified 18 chapter leaders and requested informational interviews or recommendations for interviewees. This snowball sampling procedure was supplemented by LinkedIn searches of related physical therapy professionals including physical therapists, clinic owners, technology influencers, policy makers, and leadership. We performed the study using the Standards For Reporting

Qualitative Research (SRQR) reporting guideline.

3.3.2 Data Collection

We developed a 30-45 minute semi-structured interview guide based on an initial pilot study conducted in May 2020 with 20 physical therapists from the California Bay Area through phone interviews or web-conferencing. The interview protocol was iteratively updated to explore questions about potential barriers, facilitators, and technological needs for implementing remote physical rehabilitation with telehealth. This led to the final box interview template in Figure 3.1. Two researchers then conducted interviews during July 07, 2020 - August 18, 2020. Communication was achieved through face-to-face video conferencing with the Zoom platform.



1. Tell us about your background in relation to physical therapy.
    a. Intervention and Specialties?
    b. Success Metrics for Patients?
    c. Biggest Challenges in treatment?
    d. Key influencers on practice?
    e. Session Costs?
2. How has your practice been doing over the past year? What range of technologies have you been using?
    a. Impact of COVID19?
    b. Learning about new equipment?
    c. Telehealth Solutions?
    d. Get, keep, and grow patients?
    e. Clinic financial structure?
    f. Continued care model for patients?
3. What are your current thoughts on telehealth physical therapy?
    a. Likes?
    b. Dislikes?
    c. Patient Involvement?
    d. Anything you'd like to see changed with the technology?
4. What if there was a 3D virtual environment for PTs and patients to meet? Like being in the same room without being able to physically touch each other? Any thoughts?
    a. Usability Concerns?
    b. Success Metrics?
    c. Patient Involvement?
    d. If anything was possible, what would you want to see?
5. Are there any questions we should've asked you?
6. Do you know other physical therapists from different facilities that would be willing to talk with us?

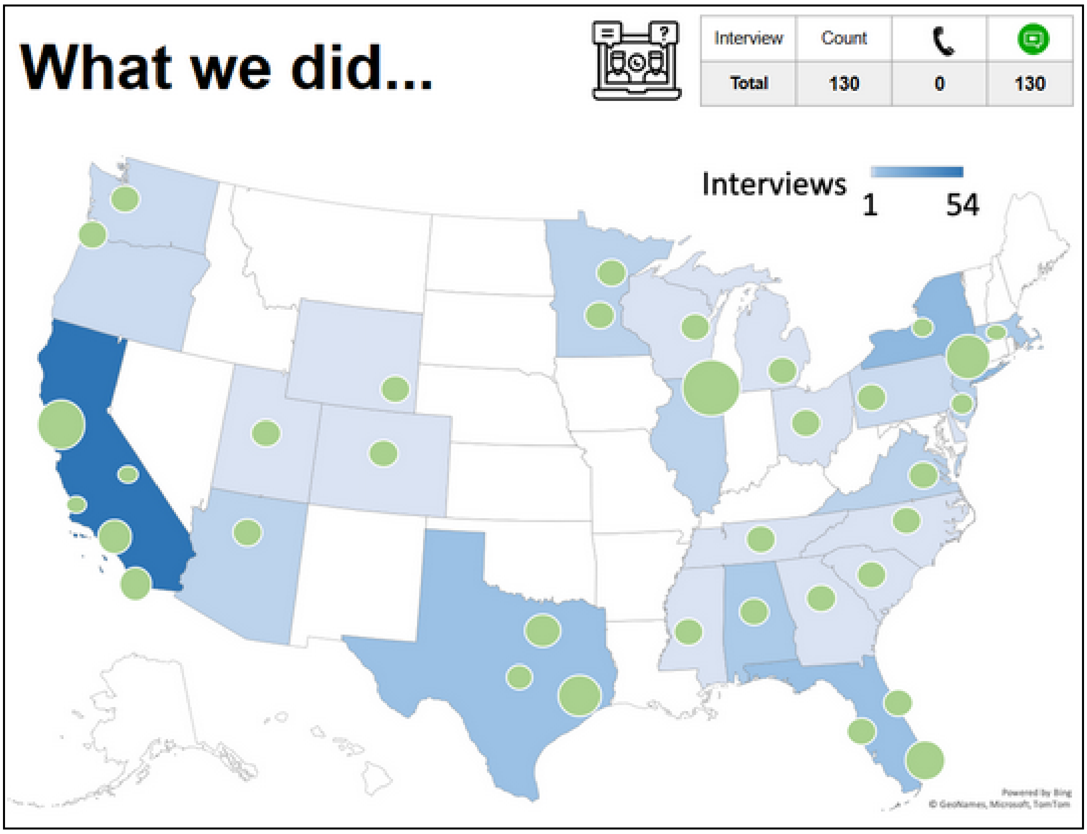
Figure 3.1: Box interview question guide for physical rehabilitation professionals.

3.3.3 Data Analysis

Interviewers took iterative notes and labeled emergent themes as the interviews progressed. Data was analyzed in a procedure where researchers identified interview themes and codes using constant comparative analysis [?] as an ongoing measure. Through this process, we identified a range of telehealth values, barriers, and needs based on participant experiences. An initial codebook categorized telehealth physical therapy interventions by value proposition, patient-therapist relationship, therapy channels, and clinical costs.

3.4 Results

We conducted 130 in-depth interviews in face-to-face live video interviews across 26 states, with most interviews conducted in Northern California (38), Southern California (16), New York (11), Florida (6), and Texas (6) as shown in Figure 3.2. Interviewees identified as physical therapists (96) with a subset identifying as occupational therapists (4), physical therapy assistants/technicians (5), or physical rehabilitation management experts (25). Participants included 74 doctors of physical therapy with specializations in outpatient (27), sports (18), orthopedic (16), home health (14), neurological (13), travel (4), and geriatrics (2). Additionally, 40 interviewees were clinic owners, 55 had experience in care management roles,50 considered themselves to be technology influencers in their workplace, 15 self-identified as being a patient of physical rehabilitation during the pandemic, and 18 were actively involved in leadership or rehabilitation policy. Interviewee healthcare institutions ranged from large hospitals to small independent clinics.



What we did...

Interview	Count	☎	💬
Total	130	0	130

Figure 3.2: Remote interview locations with 130 physical rehabilitation professionals.

The interviews identified key trends in telehealth for physical rehabilitation as a result of the pandemic. Physical Therapy typically consists of the patient going to the clinic 1-2 times per week for several weeks, months, or even years. The patient is often given exercises to do at home, which studies show have low compliance rates [**?**, **?**]. After in-person treatment, depending on the type of insurance, the patient is discharged from clinical therapy and advised to continue exercises at home to further the recovery. Most fail to continue with at-home exercises due to lack of motivation, repetitive exercises, and lack of accountability. In many areas, there are not enough therapists to meet the demand of all surrounding patients leading to difficult scheduling and less frequent clinic visits.

With shelter in place orders across the nation, clinics responded with a range of options: discontinued care for patients altogether, implemented social distancing, or adopted teletherapy using video conferencing. These responses impacted low income and rural communities especially in medical deserts with fewer health professionals relative to more affluent areas. Video conferencing enables communication, but lacks interpersonal connections between physician and patient that stems from interaction. Additionally, the physician is unable to obtain functional health metrics that help monitor progression of conditions such as heart rate or joint angles and forces for physical movement.

3.4.1 Key Findings

3.4.1.1 Physical Therapy with Telehealth, During and After the COVID-19 Pandemic

A recurring theme was the severe concern of integrating telehealth for physical therapy both during and after the COVID-19 pandemic. Many physical therapists and outpatient rehabilitation clinic owners perceived telehealth adoption to be accelerated by the impact of the pandemic, noting that telehealth for physical therapy is here to stay. Therapists explained their biggest concerns with telehealth was the lack of hands-on evaluation (e.g. manual muscle testing) and obtaining evidence-based measures of recovery such as goniometer measurements for range of motion. Most clinics turned to remote video conferencing solutions to facilitate physical therapy; however, therapists shared frustration with 2D video displays: an inability to properly observe evaluative tests such as berg balance or time-up-and-go due to limited space in the patient's home, technical difficulties in setting up webcams, and being unable to tilt the camera around the patient to view their movements as needed. Therapists shared that "all [video] telehealth can do is functional testing. I need to get a range of motion and strength assessments" with many PTs feeling that they have become "virtual life coaches instead of physical therapists" due to the limitations of phone-based or video physical therapy. Another large concern was the uncertainty of Medicare coverage in telehealth, where many clinics providing Medicare services felt indecision in investing towards telehealth without a guarantee that treatment will be reimbursed without pay cuts.

On the other hand, many interviewees noted benefits of incorporating physical therapy through telehealth. One clinical director shared "Telehealth is like [the food delivery app] Uber eats, you can reinforce it as it's convenient and patients know it exists now. They've gotten used to the convenience of telehealth and it's really here to stay thanks to COVID." Other physical therapists noted that patient no-shows were reduced, follow-up appointment retention increased, and patient visit throughput for therapists was increased due to travel time being removed and greater autonomy with online scheduling and web conference systems.

3.4.1.2 From Rehab to Prehab, Shifts in Physical Therapist Goals and Telehealth from the Pandemic

Interviewees shared many perspectives about insurance and concierge based physical therapy. Key pain points regarding insurance were noted: uncertainty of future medicare policies, perceived lowering standards of treatment due to insurance evidence requirements, and turning clinics into "patient mills" where therapists per patient are increased with minimal visit time to get enough revenue to keep their clinics afloat. Patient mills were a recurring concern between in-patient and out-patient physical therapists where pay cuts in insurance appeared to be one of the largest causes amplified by the global pandemic. Consequently, clinical directors and traveling therapists conveyed that out-of-network coverage is a growing trend where physical therapists who are suffering from burnout and insurance are turning instead towards concierge based physical therapy. Another self-practicing physical therapist shared “every physical therapist that opens a clinic dreams of getting their patients into fully cash-based, insurance is a means to an end to get more patients, but the dream is fully cash-based to give us the freedom to help our patients as they need it.”

Thus, a majority of interviewees, self-practicing or clinical owning physical therapists, are actively searching for new business models to support preventative treatment, with cash-based or concierge physical therapy as an avenue to support this model. Interviewees argued that insurance based physical therapy is largely a rehabilitative market, reacting to injury without emphasizing or recognizing the need of preventive rehabilitation (prehab). One therapist shared “physical therapy should last years, become a lifestyle [for the patient], but in reality the minimum is done for patients who often fall back into cycles of treatment.” This perspective was amplified by the rapid adoption of telehealth, especially for continued care. Many were laid off at the start of the pandemic due to clinics closing due to decreased patient visits or lack of infrastructure to adopt telehealth as a business model. A majority of therapists in this situation turned towards self-practicing concierge services and preventative care, effectively becoming traveling physical therapists through a hybrid mix of occasional visits to patients home and conducting check-ins with video conferencing. Many self-practicing and clinic owning therapists believed that after the pandemic, their practices will likely continue with telehealth or hybrid models of video visits with in person. Many emphasized moving away from injury to preventative care, where follow up visits can be conducted through telehealth for a lower cost but at a higher visit rate. Telehealth was often seen as a solution to facilitating prehab through continued care: getting the tools into the hands of patients before they even become patients with video visits, asynchronous messaging between therapist-patients, and a greater rate of visits.

3.4.1.3 Speculative technology solutions and future directions for telehealth

Some clinics took the opportunity to incorporate high-tech solutions into facilitating telehealth: this included smartphone based patient videos for exercises, wearable sensors with IMUs to track range of motion, and exoskeletons for at-home supported movements and exercises. Interviewees were encouraged to speculate about possible improvements for telehealth beyond video based interaction. Therapists were asked "what if you could stand in the same room with your patient, walk around them in 3D, but be unable to touch them?" Four recurring themes were revealed: (1) enhancing therapist perception with immersive 3D technologies to effectively enable therapists and patients to be in the same room and walk around each other remotely, (2) the ability to gather evidence-based measures of patient performance remotely through patient motion capture, (3) automating documentation and measurements to give the therapist more one-on-one time with the patient, and (4) provide a means for therapists to run their own practice or clinic through the cloud to reduce the overhead costs of starting their own practice or running a practice during the pandemic. These themes are shown in Figure 3.3.

3.5 Discussion

The global COVID-19 pandemic has vastly accelerated the adoption of telehealth for physical rehabilitation. Many therapists shared that this event has fast tracked the technological remodeling of modern clinics as both patients and therapists have become used to the technical needs of living in a virtual-social workplace with sheltering in place. While telehealth was initially viewed as a flawed practice for physical therapy, clinicians have found strengths in remote healthcare such as increased patient retention, reduced travel time, and less operating costs. The largest challenges of telehealth was the inability to collect evidence-based success measures and evaluate patient movements. Metrics of capturing range of motion, muscle strength, and balance risk were all common measures for in-person physical therapy, but these were lost in telehealth. Nonetheless, a common consensus was that telehealth in physical rehabilitation is here to stay. Beyond the pandemic, many practitioners shared they would switch to a hybrid model of in-person treatment and telehealth or entirely adopt telehealth practices to run "physical therapy clinics in the cloud." In examining future approaches for mitigating limitations in telehealth, immersive virtual environments were found to be one promising solution: enabling 3D virtual worlds for patients and therapists to meet, walk around one another, and examine exercise movements.

Telehealth was seen as a possible solution to therapist burnout and barriers to entry in starting independent clinics. While therapists pointed out limitations in video and text based communication platforms, there was growing disconnect with returning "back to normal" after the pandemic subsided. Many therapists shared that medicare based clinics can often feel like "patient mills" – treatment overloaded by patient size and conforming to insurance restrictions (e.g. medicare compliance). Many therapists that shared the desire to start their own clinics or switch to independent practitioners, but were deterred by student debt or clinic startup costs. When speculating about the future, therapists viewed the cash-based physical therapy market as a potential entry point to mitigate the startup costs without the burnout from adhering to insurance restrictions.

Telehealth's largest challenge with physical rehabilitation was the loss of hands-on evaluation tools. Many therapists never received formal training for conducting remote treatment during the pandemic, and the loss of hand-based evaluation (e.g. manual muscle testing) was often referred to as the largest challenge. Throughout the interviews, four recurring tests were often discussed: 1) Range of Motion (e.g. Goniometer), 2) Strength (e.g. manual muscle testing, or Isokinetic Dynamometers), 3) Coordination (e.g. Timed up and Go), and 4) Balance (e.g. Berg Balance Assessment). With the loss of these measures, therapists resorted to having patients self-test or became "life coaches," meeting virtually to provide advice rather than treatment. While some therapists acclimated to utilizing digital measurement tools (e.g. virtual goniometers from video), many struggled to obtain proper lighting, full depth of movements, and or camera quality to perform a visual evaluation. This informed our speculative component of the interview to determine how emerging technology could mitigate these challenges.

Immersive virtual environments offer a powerful opportunity to transform physical rehabilitation telehealth. In discussing the incorporation of 3D virtual environments for treating patients, therapists saw opportunity for capturing range of motion and strength assessments utilizing head-mounted display motion capture. The incorporation of a 3D virtual room in which patients and therapists could meet virtually through avatars seemed to be a potential solution to the existing limitations of video based telehealth. While the loss of hands-on evaluation was a barrier, therapists believed such an approach would enhance continued care, retention, and reliability of patient progression. Therapists did note many possible risks with such an environment: (1) patients with unstable balance risk injury without in person supervision, (2) virtual representation of the patient's movements must be accurate in order to reduce movement bias, and (3) motion sickness from the headset may cause discomfort.

3.5.1 Limitations

While we believe that the results we report in this paper are generalizable, there are some limitations of the study that need to be taken into account. The COVID-19 lockdown forced us to reconfigure interviewing as video conferences rather than in person. This limited the fidelity of the demonstrations we can witness from the health care professionals we interviewed. Secondly, with our sample states, our results might not be as relevant to states with different regulations or practices. Finally, the majority of our interviewees were physical therapists. Our results might differ if the professionals were different.

3.6 Conclusion

In this research on the uses of telehealth in physical rehabilitation throughout COVID-19 restrictions, our empirical study with physical therapists offers insights into new directions for telehealth after the pandemic. Common themes emerged regarding potential solutions for working around issues of presence or measurement data collection. These include: (1) immersive 3D technologies which allow PTs and clients to walk around each other remotely, (2) evidence-based measures, (3) automating documentation and measurements, and (4) provide a means for therapists to run their own practice or clinic through the cloud. These approaches may enable PTs to reduce the overhead costs of starting or maintaining practice in event that future pandemics require limited in person care or therapeutic treatment. In summary, the knowledge gained from these interviews indicate significant opportunities to expand the capabilities and reach of physical therapists for rehabilitation and preventative care via telehealth.

3.7 Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No #2037917. Such work was also supported by the 2020 Seed Fund Award 2020-0000000044 from CITRIS and the Banatao Institute at the University of California, a statewide research institute operated by the University of California to facilitate the real-world application of technological research. Additionally, Elor was supported by the University of California, Santa Cruz Global Community Health Well-beings 2020 Fellows Program and ARCS Foundation 2020 Fellowship. We also thanks Dana Seth, DPT, of the Austin Physical Therapy Clinics, who provided assistance and feedback during the early stages of the study design. Additionally, Greg Gewickey and Ash Robbins provided feedback throughout the study and recommended key individuals and stakeholders to interview.

Category	Challenges and Opportunities
Risks associated with telehealth for physical rehabilitation	1. Remote interaction is the biggest pain point. a. Inability to collect evidence-based data: Loss of therapist hands on interaction for manual muscle testing and physical assessment. b. A lack of measurement tools for accurate range of motion or muscle strength metrics. 2. Technology driven interventions a. Patient-therapist education with telehealth user interfaces (i.e. a need for user friendly interfaces for non-tech savvy patients).. b. Service limitations (i.e. internet connection, computer hardware latency, webcam quality, physical location to perform physical rehabilitation in) 3. Patient interaction a. Patient safety (therapist is unable to be there with the patient to ensure they don't fall or hurt themselves). b. Loss of depth information (2D videos provide visualization challenges for viewing patient movement and guiding patients through movement as a therapist) 4. Policy and Maintenance a. Therapists are unsure if telehealth will continue to be covered by insurance/medicare. b. Therapists are challenged in building virtual reputation.
Benefits associated with telehealth for physical rehabilitation	1. Convenience: Therapists and patients are no longer limited by physical location or travel times to receive care. a. Patients are more likely to be retained by therapists for continued care. b. Therapists are able to reach more communities. c. Reduces the amount of patient "no-shows." 2. Efficiency: Therapists can reduce brick and mortar physical costs through the "virtual clinic." a. Remove the need to own a physical office to enable physical therapy treatment. b. Automate scheduling and documentation through telehealth integrated with EMS platforms.
Value Propositions of Telehealth and Virtual Reality (i.e. how might emerging technologies influence telehealth)	1. Predictive analytics for providing quantitative patient based monitoring and assessment. [N=127] 2. Interactive 3D environments for facilitating patient-therapist interaction. [N=114] 3. Enabling physical therapy practice through virtual platforms to reduce clinical/brick-and-mortar costs [N=104] 4. Facilitating patient continued care through virtual platforms and gamification [N=87]. 5. Facilitating automated patient documentation and measurements [N=38] 6. Providing patient to therapists recruitment through virtual platforms [N=28] 7. Ensuring user friendly virtual platform interfaces for telehealth [N=24]

Figure 3.3: Emerging themes from interviews with 130 Physical Rehabilitation Professionals.

Chapter 4

Gaming Beyond the Novelty-Effect of Immersive Virtual Reality for Physical Rehabilitation

4.1 Summary

Immersive Virtual Reality (iVR) Head-Mounted Display (HMD) systems paired with serious exercise games can positively augment physical rehabilitation process from both engagement and analytics perspectives. This paper presents a serious game for iVR HMD based long term upper-extremity exercise. We demonstrate the capabilities of our game through a case study with five users recovering from upper-extremity injuries. We examine how our program maintains engagement and motivation over eight weeks, where users completed bi-weekly prescribed movements framed as protecting a virtual butterfly. We assess user experiences through a mixture of biomarkers from brainwave, heart rate, and galvanic skin response recorded at runtime as well as motion capture and behavioral game data. Our results suggest that the iVR game was an effective medium

in inducing high compliance, physical performance, and biometric changes even with increasing difficulty beyond the novelty effect period. We conclude with considerations of future work for iVR physical therapy games that adapt to biometric response.

4.2 Introduction

With the mass commercial adoption of immersive Virtual Reality (iVR) based Head-Mounted Display systems (HMDs) and over 200 million headsets sold since 2016, iVR systems have almost become a common household gadget [52]. On a parallel track, rehabilitation research, including physical and cognitive work that incorporates VR based interventions, has been on the rise due to the ability to create programmable immersive experiences that can directly influence human behavior [87, 86, 69]. The ability to run conventional therapy in a virtual environment can be paired with high-fidelity motion capture, telepresence capabilities, and accessible experiences [72, 419]. Through serious games, immersive environments with commercial HMDs can be programmed to translate therapeutic goals into game mechanics, making the therapies more enjoyable [420, 421, 422, 12, 13, 227].

It has been shown that iVR can be successfully used for treating Post Traumatic Stress Disorder [64], Borderline Personality Disorder [65], various phobias [66, 69], schizophrenia [68], and others. The detachment from reality and immersion in a virtual world can reduce discomfort, even as far as minimizing pain when compared to clinical analgesic treatments [61]. Strong immersive stimuli through a VR system together with the ability to combine presence and emotion in a virtual world is key to influencing user behavior [64]. However, quantifying this success is often difficult due to system constraints and a lack of computational power [33].

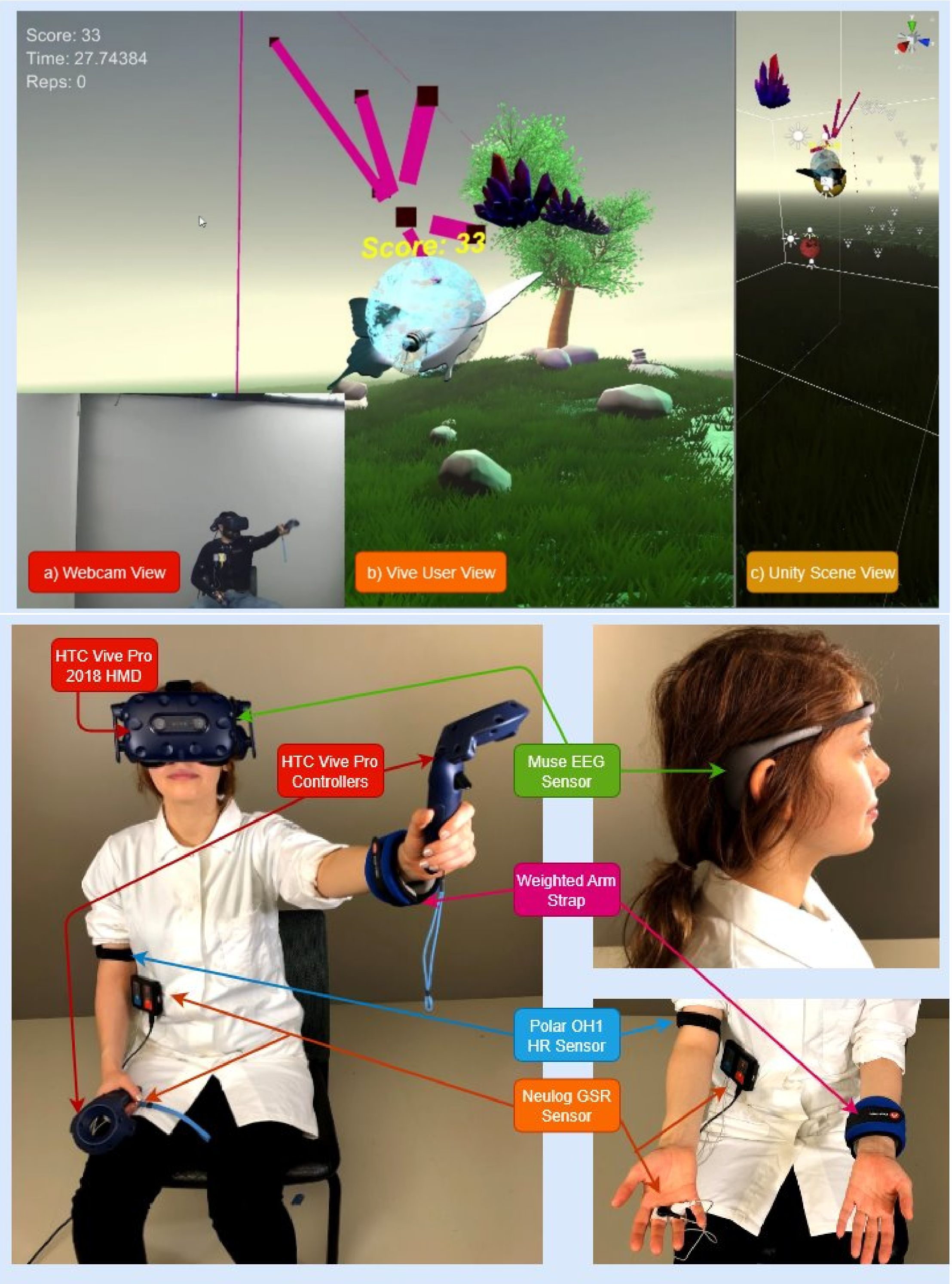


Figure 4.1: A user plays Project Butterfly during their bi-weekly exercise session: a) The user physically matches the butterfly's position to protect it while also catching crystals to infer the motion path. b) The user's view of the game through the Vive HMD. Score is visible on the user's bubble shield, with time and repetitions visible for the evaluator interface. c) The scene view of the user in unity. The bottom half of the figure depicts sensor peripherals utilized during Project Butterfly user testing.

Similarly to iVR, research in biometric sensing has seen explosive growth over the past decade. We argue that biofeedback may serve as a versatile tool to quantify the success of an iVR based physical therapy experience. Brain-Computer Interfaces incorporating Electroencephalogram (EEG) devices have become more affordable and user-friendly – they come with computational techniques for understanding user engagement and intent in the fields of medical, entertainment, education, gaming, and more [142]. The analysis of different brainwave frequencies has been correlated to different psychological functions, such as the 8-13 Hz Alpha band relating to stress [150], the 13-32 Hz Beta band relating to focus [151, 152], the 0.5-4 Hz Delta band relating to awareness [155, 156], the 4-8 Hz Theta band relating to sensorimotor processing [159, 160], and the Gamma band of 32-100 Hz related to cognition [161, 163].

Researchers are combining these interfaces with other forms of multi-modal biometric data collection such as Galvanic Skin Response (GSR) and Heart Rate (HR) to increase affective inference. Through measuring the changes in skin resistance, GSR has been linked to emotional arousal [135]. Quantifying the intensity of GSR may enable researchers to record emotional response with stimuli such as music, television, and gaming [137, 136]. Additionally, it has been shown that the accuracy of the non-invasive prediction of the psychological response can be increased by using data fusion algorithms that integrate GSR, HR and EEG readings [143, 140, 138, 133]. These biometric measures paired with iVR therapy may yield much potential for understanding physical exercise experiences, but are not often incorporated in the traditional physical therapy regimen.

Traditional outpatient physical therapy usually involves clinical visits where patients perform exercises, receive evaluations, and undergo in-person manual therapy interventions. Better clinical outcomes are often associated with a well established schedule of at-home exercises following the clinical visits [423]. Many meta-analyses and reviews have suggested that virtual reality-based rehabilitation techniques, some with biometric sensing [424, 425, 426], could outperform traditional physical therapy [427, 428, 429]. Some of these studies have shown that iVR therapy can be effective for weeks or months [430, 431, 432]. However, although there is mounting evidence that VR HMD-based rehabilitation systems could benefit from integration with with biometric, physical, and user-reported analysis, there is a lack of studies with extended period investigating this phenomenon.

The study reported in this paper aims to answer the question of: 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 Doctoral Physical Therapists, with 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" [433, 434]. In this study, we extend these principles to stimulate 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.

In reflecting on Rego et al.'s taxonomy of serious games for rehabilitation, we present a two-month user study on a motor game that utilizes motion-tracking for a head-mounted display based 3D single-player action experience with adaptability, progress monitoring, performance feedback, and clinic-to-home portability [435]. We define serious games as games whose purposes are beyond entertainment only [435]. We define iVR-based serious games as immersive experiences that incorporate 3D user interaction and motion capture, often with an HMD. To the best of our knowledge, this study is one of the first to leverage an immersive VR HMD based serious game for an extended physical therapy period and that examines multi-modal biometric feedback and physical performance.

4.3 System Design

Our physical therapy system is based on a game called "Project Butterfly" (PBF) that was proposed by Elor et al [234, 436]. Previously, PBF explored the feasibility of a virtual reality enhanced exo-skeleton for post-stroke and elderly assistance through two exercises, but was not designed or tested for upper extremity physical therapy over a extended period of time with varying custom exercise movements as reported in this study. The game was built using the Unity 2018.2.11f1 Game Engine with SteamVR and incorporates the HTC Vive Pro 2018 (Vive). Vive uses outside-in tracking through a constellation of “lighthouse” laser systems for pose collection in a 3D 4x4m space [52]. It has been verified in previous studies to analyze therapeutic gamification [300, 301], postural analysis [258], and accuracy for research data collection [259].

To explore the effects of extended upper-limb rehabilitation, we heavily modified PBF to create a new gaming experience. These improvements were designed and suggested by the collaborating therapists to make the game customizable and aid best in the rehabilitation process. The updated goal of the PBF is to safeguard a virtually flying butterfly from adverse weather and flying projectiles using a translucent protective “bubble shield” that is controlled by the player through the Vive Hand Controller. The flying pattern of the butterfly recreates the rehabilitation exercises, and the game was modified to allow dynamic adjustment of movements through therapist-led motion capture, enabling the dynamic control of pace and position for the prescribed exercise. Specifically, using the HTC Vive Controllers in PBF, therapists can record motion paths by entering a custom unity scene and performing exercises by holding the controller trigger while performing the desired movement. Movement paths are remotely uploaded into the unity environment through comma-separated value format. The 3D motion vector path is then previewed, confirmed, and named by the therapist, which is then adjusted to user arm length during gameplay. The user is required to follow the path of the butterfly with a 0.1 meters error, where they are awarded a score point (accompanied by audio and haptic feedback) for each half a second they successfully protect the butterfly. This includes a scoring system for the user to engage in self-competition.

Gameplay sessions were updated to provide feedback to both the user and the therapist by collecting EEG, GSR, HR, motion capture, and player behavior datum. Each exercise is prescribed by our collaborating physical therapists to establish range of motion in the first stage of rehabilitation and complex movement in the last stage. The user starts following the path by activating an animation with the controller placed on the stationary butterfly for three seconds at the start of the game. Collectible game projectiles were added in the shape of crystals, and spawn from a distance while veering towards the butterfly as a motion path indicator for the user to follow their exercise. The gameplay of the new PBF version is shown in Figure 4.1.

Additionally, the updated project applies concepts from Self-Determination Theory (SDT [437]) to intrinsically motivate users to become the guardian of a butterfly (relatedness), gain mastery of controlling a virtual shield in following the butterfly's path (competence), and be in full control of their upper extremity movement with progressive difficulty (autonomy). Previous studies have examined motivation in exergames through gamification (challenge and score) [438], self-competition and virtual trainers [439], music [440], sensory feedback [441], rhythmic clapping [442]. PBF differs from these exergames in a gameplay mechanic aimed towards relatedness: enacting a helping behavior on the user to protect the butterfly, building a gradual sense of attachment and belonging to the butterfly through becoming it's guardian. Research has shown that helping behavior can provide a variety of material, social, and self-rewards such as mood enhancement to the helper [443]. Subsequently, the care-helping relationship was found to be strong mediator for stimulating planned, long-term helping behaviors in user participation [444]. Through iVR, we extend these concepts into an immersive environment. Given that iVR therapeutic intervention's success is often attributed to the influence of immersion on users in terms of enhancing the relationship between presence and emotion [64, 69], the system proposed in this paper provides a mixture of visual, audio, and haptic feedback for upper body movements helping the user to overcome the adversity of physical task-based objectives.

4.4 User Study

The methods used in the study reported in this paper received Institutional Review Board's approval from the University of California - Santa Cruz Office of Research Compliance Administration. Five college students recovering from upper limb injuries (including shoulder dislocation, shoulder impingement, and rotator cuff tears) consented to volunteer. These participants consisted of one female and four males ranging from 21 to 28 years old. Additionally, the participants were no longer performing in-person physical therapies, nor continuing at-home therapy exercises, thereby mimicking our target population. These users agreed to volunteer for two weekly sessions for eight weeks, although, as the consent form stated, they were free to drop out of the study at any time. Participants were offered a $50 USD gift card upon completing the study (or were paid about $3.12 USD per session). All users completed all sessions of the study.

4.4.1 Data Collection

Figure 4.1 shows the experimental setup. Through Vive and the Unity Game Engine, motion capture and behavior game data were recorded and stored during runtime at 90 Hz using a data exportation method developed in previous studies by Elor et al. [234, 13]. Motions were captured through utilizing the HTC Vive Pro 2018's outside-in tracking algorithm by a constellation of "lighthouse" laser systems [52]. Runtime recording of brainwave (EEG), GSR, and HR data was also performed to complement the Vive data. EEG was sensed from the pre-frontal cortex (TP9/10, AF1/7 locations) and recorded using InteraXon Muse 2, a commercially available brain sensing headband [170]. While Muse is a relatively new EEG commercial device, it has successfully been used in other studies to infer mental state, analyze event potentials, and record biofeedback [393, 394, 392]. For GSR, we utilized the Neulog GSR logger sensor NUL-217 [390] to measure the skin's conductivity between the fingers of the non-dominant hand. HR was recorded through the LED optical sensor Polar OH1 [389]. Every sensor was chosen with accessibility and cost as a factor. These devices are easy to set up for a user at home, with dry contact after wiping with saline to reduce the costs of any EEG gels or sticktrodes for HR and GSR that are usually required when using many clinical-grade sensors. Sensor locations and equipment setup used for each user testing session are seen in Figure 4.1.

At the end of each testing session, users were asked a series of Likert scale questions regarding their rehabilitation session experience with PBF. Survey questionnaires were inspired from a Jennett et al. survey, which measures immersion in games [397]. The survey was implemented to focus on behavioral engagement and a self-reported emotional response between multiple exercise sessions over many weeks. Reflecting on Doherty and Doherty's review of engagement in human-computer interaction, we define engagement assessment as a behavioral understanding of whether a user desires to and can effectively use a system [445]. For understanding emotional response, we utilized the circumplex model of affect to assess self-reported emotions from users between exercise sessions [446, 447]. The goals of these surveys were to understand if users would self-perceive PBF as engaging, immersive, and positive throughout the two months as the users become acclimated to the game during rehabilitation. An exit interview was also conducted at the end of the two months to better understand participants' subjective feedback.

4.4.2 Protocol

The study followed each of the five users during their prescribed rehabilitative exercises through two protocol phases: Foundation and Challenge. The Foundation Protocol instigated basic motion primitives that emphasized recovery of Range of Motion (ROM) and shoulder strength through three simple exercises: Forward Arm Raise [FAR], Side Arm Raise [SAR], and Shoulder Rotation [SR]. The goal of the Foundation Protocol was to acclimate the user into performing physical therapy in iVR with minimal weight resistance to reduce the possibility of injury as recommended by our collaborating therapists, and see if the user could maintain compliance over an extended period of iVR exposure. The Challenge Protocol added four complex motions that further pushed ROM recovery while including the original Foundation Protocol's exercises using increasingly higher weights. The goal of the Challenge Protocol was to challenge the user with continuing weight increases and complex movements to investigate the effects of more complex exercises on compliance and engagement. The complex movement exercises added were External Rotation [EXR], Abducted Rotation [ABR], Mixed Press [MXDPR], and Mixed Circles [MXDPC]. These movements and their protocol inclusions can be seen in Figure 4.2.

Each protocol consisted of five sessions, excluding the initial tutorial session. Two evaluators were always present to monitor the user for irregular activity and record qualitative observations. Exercises were played in circuit rotation and can be seen in Figure 4.2. Weighted straps were used for the motions FAR, SAR, and SR and the weights were gradually increased throughout the study to match the participant's capability. The protocol for increasing weight was determined with our collaborating therapists to minimize the chance of injury. Participants usually started at 1-2 pounds and weights were increased by 1-1.5 pounds after comfortably completing two rounds of gameplay. The order of the movement was kept consistent to monitor biometric results for examination of data between sessions, as illustrated in Figure 4.2.

Users performed three sets of the exercises for the Foundation Protocol and two sets of the seven exercises during the Challenge Protocol. A break of approximately 90 seconds was given between each exercise. The user then performed 60 seconds of the prescribed movements with biofeedback, game data, and video recorded. The entire sessions lasted forty-five minutes to an hour, including post-user surveys. For our analysis, we examined the change from the resting baseline to the gameplay data. A biometric baseline was recorded at the beginning of each session where the user wore all of the sensors shown in Figure 4.1 while sitting in a relaxed state with the main screen of PBF displayed on the Vive. The first session, dubbed session “zero”, of each protocol phase consisted of a basic range of motion test and tutorial gameplay where the participant performed all new exercises of the protocol. By playing the games on this initial session, we hoped to limit the novelty effect. It should be noted that data was not collected on these tutorial sessions, as the protocol was unique to each individual to answer questions and help with acclimation to the iVR environment.

4.5 Results

Results from session data were post-processed using the Mathworks Matlab 2018b Statistics and Machine Learning Toolbox [400]. 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 4.3. 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 4.4 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.

For brainwave response, neural activities were measures at all sessions and all protocols, as shown in Figure 4.5. 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 [170] 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 4.6. While playing PBF, users in the Foundation Protocol tended to blink less than their resting state for every session (with 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 4.7 (engagement based) and Figure 4.8 (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 4.7. Similarly, the majority of users reported a positive range of emotions for each exercise ranging from Happy/Joyful, Excited/Motivated, and Relaxed, as shown in Figure 4.8. 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.

4.6 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 [150], focus [151, 152], awareness [155, 156], motor [159, 160], and cognition [161, 163]) 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 [448]. 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 [449]. The survey responses also showed that users felt engaged by each protocol. In Q1 (Figure 4.7), 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.

4.7 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 runtime biofeedback and adaptive difficulty with 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 runtime may yield immense potential. Subsequently, there are far more butterflies to follow on the road ahead.

4.8 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 participants of this study, without their time and effort, none of this would be possible.

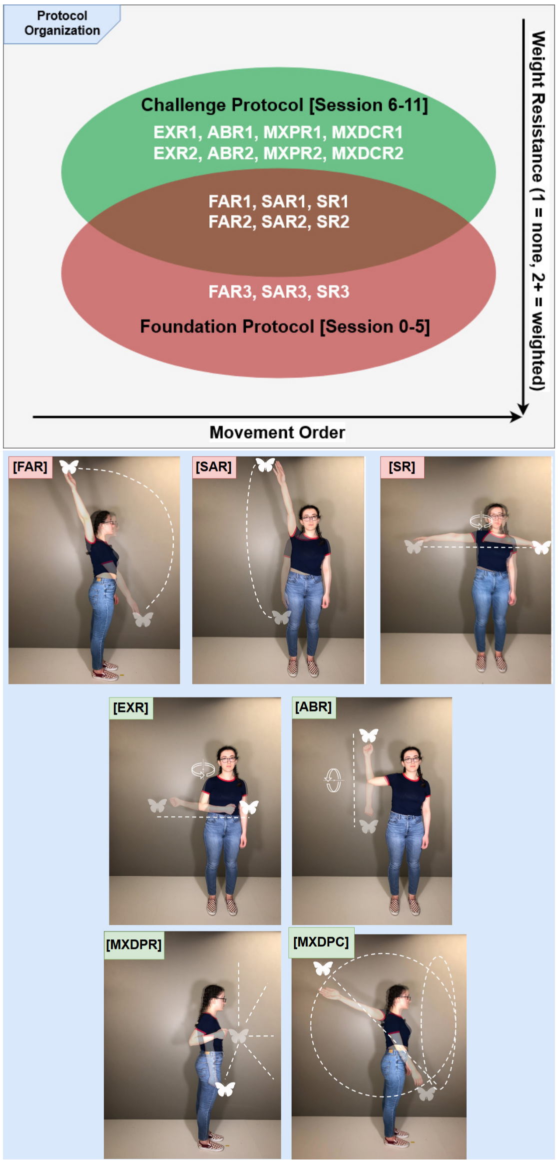


Figure 4.2: Movements tested in the Foundation and Challenge protocols. The dotted line shows the butterfly's path, while the arrows indicate the axis of rotation for relevant movements.

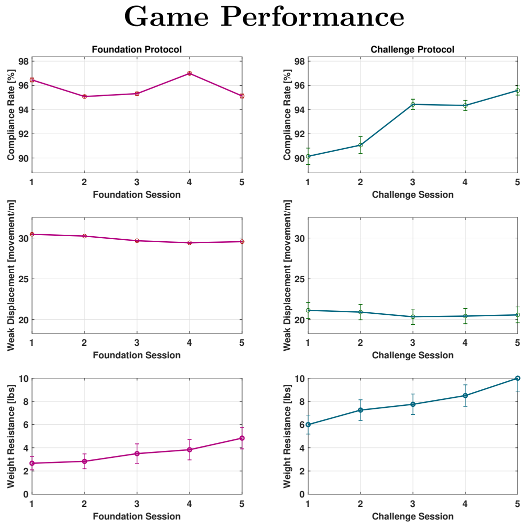


Figure 4.3: Game performance between Foundation Protocol (in red of 225 recorded exercises) and Challenge Protocol (in green of 350 recorded exercises). Row one shows compliance, where compliance is defined as the total time protecting the butterfly over the game's total time. Row two shows the mean 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).

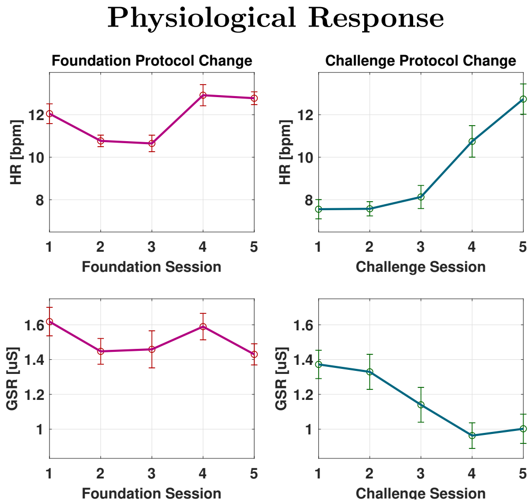


Figure 4.4: Physiological HR and GSR responses from gameplay are shown. Row one illustrates mean change from resting state of heart rate. Row two illustrates mean change from resting state of galvanic skin response. Biometric change is calculated as the offset between gameplay biometrics against resting-state biometrics. Error bars indicate standard error.

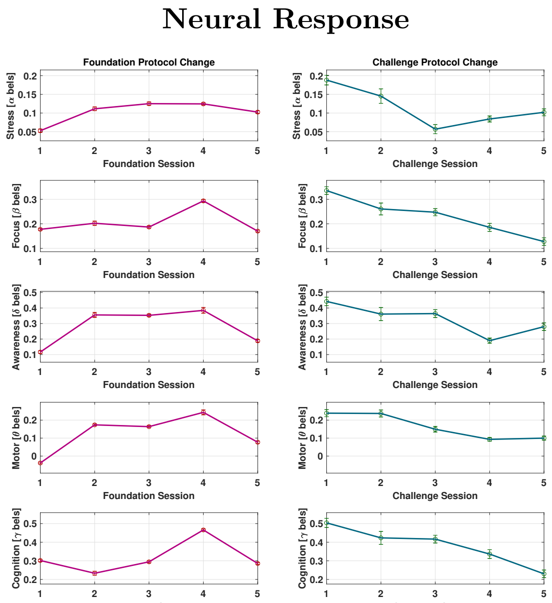


Figure 4.5: EEG responses between Foundation Protocol (in red of 225 recorded exercises) and Challenge Protocol (in green of 350 recorded exercises). Rows 1-5 show Alpha, Beta, Delta, Theta, and Gamma bands resting state change respectively. Error bars indicate standard error.

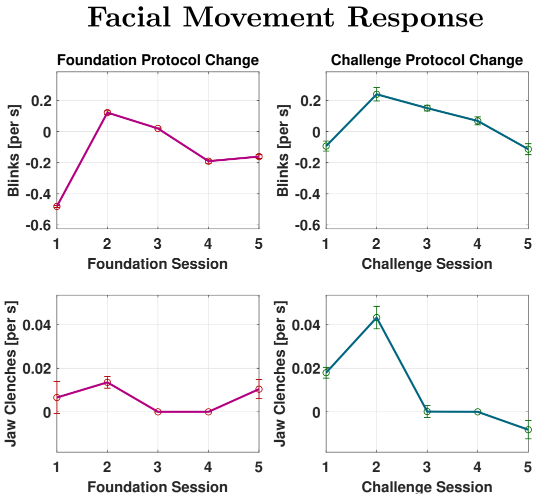


Figure 4.6: Facial muscle movements recorded with Muse between Foundation Protocol (in red of 225 recorded exercises) and Challenge Protocol (in green of 350 recorded exercises). Row one shows the mean resting state change of blinks per second. Row two shows the mean change of jaw clenches per second from resting state.

Engagement Survey Response

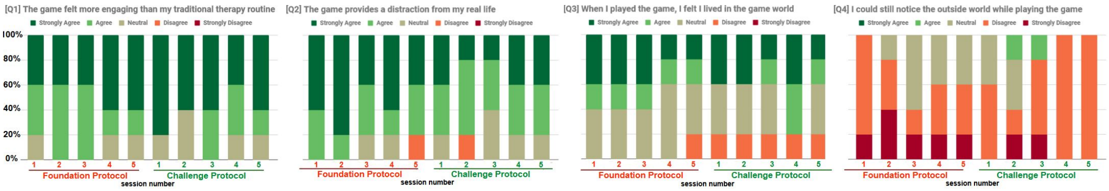


Figure 4.7: Survey responses on engagement from 5 subjects, with 1=strongly disagree and 5=strongly agree.

Emotion Survey Responses

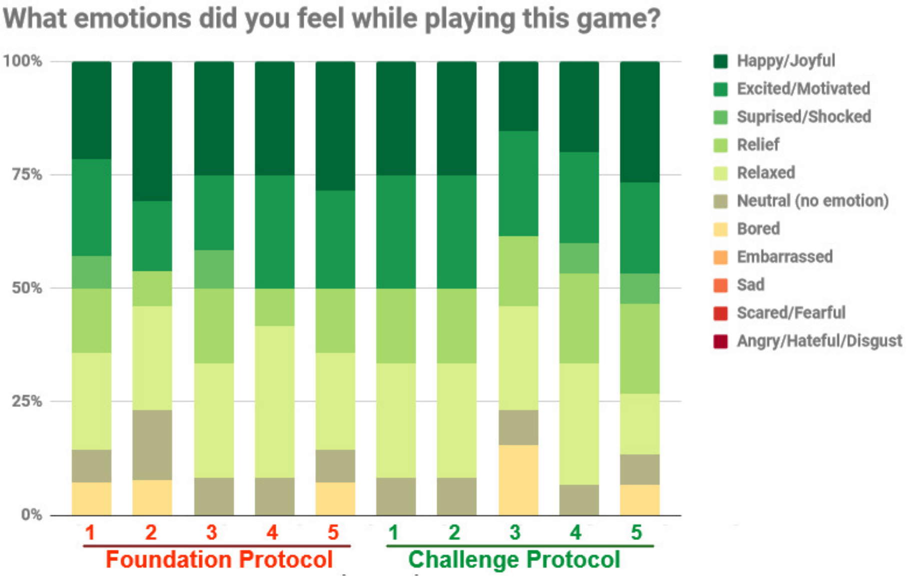


Figure 4.8: The self reported emotions ratios felt by users from post-gameplay survey.

Part VI

CONCLUSIONS

Chapter 1

Reflection

The work presented in this dissertation explores the design and evaluation of iVR experiences in physical exercise gaming, personalizing systems for emotional intelligence, utilizing systems for physical intelligence, and adapting iVR serious games for physical rehabilitation from choosing the right iVR medium to extended usage in rehabilitation protocol. Physical rehabilitation is extensively challenged in providing accessible, affordable, and engaging experiences from both a telehealth and user adherence perspective. From our work, we believe that iVR can be adapted to address these issues to present serious games that stimulate user engagement to increase adherence and incorporate an intelligent biomechanical analysis of user motion capture to capture key success metrics (e.g. range of motion, joint torques, exercise compliance) for users and therapists alike. This dissertation research produced numerous conference and journal articles that all examined the design and evaluation of novel iVR systems for a variety of rehabilitation and other serious healthcare applications [12, 13, 14, 15, 2, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28]. Subsequently, this part concludes with a key summary this dissertation work and considerations for future research in immersive virtual reality for physical rehabilitation.

1.1 Key Summary

This dissertation work began by examining how two beneficial forms of physical rehabilitation could be translated into effective serious iVR games. Through Project Star Catcher [13], we found that iterative co-design with stroke survivors and adults with developmental disability were incredibly valuable in determining game domain, mechanics, difficulty, data log-file requirements. The game was tested with multiple users from Hope Services Santa Cruz and the Cabrillo College Stroke & Disability Learning Center where we found that translating Constraint Inducement Movement Therapy from physical constraint (traditionally binding the non-hemiparetic arm) to psychological incentive (affording movement in both arms but providing a score incentive to the hemiparetic arm) could help increase exercise compliance by 40%. Additionally, we were able to improve the average compliance rate by 13% (a change from 65% to 73%) by utilizing a pilot study to refine game mechanics and ensure that our stakeholders could inform the system's design [13]. We also found that many of the traditional therapy analysis methods (usually written documentation, video coding, physical analog measurement tools like goniometers) could be automated with iVR motion capture to expedite user experience analysis and produce runtime analysis metrics (e.g. weak arm usage, limb range of motion, game performance). The exercise game was also compared between a research-grade room scale system (the CAVE) against an off-the-shelf iVR HMD (the HTC Vive) [14] for users with mixed abilities while varying exercise weights and biometric recording. We found that a modern HMD such as the HTC Vive was more engaging and produced better physical exercise performance than the more expensive room-scale CAVE platform. This study, along with our bridging review of immersive media for physical rehabilitation [19], helped inform our decision to utilize commercial off-the-shelf iVR HMDs. With this pilot work, we moved on to examine how iVR exercise games could further be utilized to assist in physical rehabilitation.

Through Project Butterfly [15, 16, 22], we began directly collaborating with physical therapists to translate how rehabilitation exercises could be mapped and prescribed remotely both with and without robotic assistance. The game was iteratively tested with stakeholders from Hope Services Santa Cruz, Elderday Adult Health Care, and Dominican Healthcare to design exercise recording mechanics and a motion capture analysis pipeline for range of motion and exercise repetitions as well as the domain setting like Project Star Catcher. The goal of Project Butterfly was to create a controlled immersive media environment for the adaptable and translatable therapeutic movement, which incorporates runtime data feedback on player movement performance and behavioral analysis. It aimed to bridge the gap between therapists and researchers with users undergoing repetitive exercise and physical therapy through mapping movement by motion capture to gamified scenarios such as protecting a virtual butterfly and catching crystals. Expanding upon the work of Project Star Catcher, this experience translates Mirror Visual Feedback Therapy into an immersive virtual reality environment that requires users to protect a virtual butterfly. The game was designed over three iterations in the course of two years informed by user testing:

- Iteration One: A user protects a butterfly from heavy rain by covering the butterfly with an umbrella. Movements are pre-scripted to follow basic motion primitives for bicep curls, shoulder rotations, forward arm raise, and lateral arm raise. This iteration was tested and iteratively designed through weekly focus groups with

Hope Services California over the course of one month.

- Iteration Two: A user protects a butterfly from heavy rain by encasing the butterfly within a spherical orb. Movements are pre-scripted to follow the same basic motion primitives as iteration one. A logfile data collection was implemented using Microsoft .NET I/O Framework in C# to stream player pose, butterfly pose, time, and score to a CSV file during runtime gameplay from the Unity Engine. An evaluator interface was created to set the speed of butterfly, length of the user's arm, username, and sample rate of logfile data collection. This iteration was tested with Hope Services, Elderday Retirement Home, and Cabrillo College Stroke and Disability Learning Center. Additionally, a pilot study was run with the CRUX tensegrity exosuit to assist users with limited movement in gameplay [15].

- Iteration Three: The updated game applies concepts from Self-Determination Theory (SDT [437]) to intrinsically motivate users to become the guardian of a butterfly (relatedness), gain mastery of controlling a virtual shield in following the butterfly's path (competence), and be in full control of their upper extremity movement with progressive difficulty (autonomy). PBF differs from previous exergames in a gameplay mechanic aimed towards relatedness: enacting a helping behavior on the user to protect the butterfly, building a gradual sense of attachment and belonging to the butterfly through becoming it's guardian. Research has shown that helping behavior can provide a variety of material, social, and self-rewards such as mood enhancement to the helper [443]. Subsequently, the care-helping relationship was found to be strong mediator for stimulating planned, long-term helping behaviors in user participation [444]. Through iVR, we extend these concepts into an immersive environment. From interviewing over 130 physical therapists during the COVID-19 pandemic, we adapted the system logfile analysis to incorporate machine learning to capture joint torques and range of motion as existing telehealth solutions were limited by inability to view users in 3D and capture these metrics. Moreover, we extended the therapist interaction such that movements can be customized and recorded before gameplay by holding down the trigger of the HTC Vive controller and performing the desired motion to prescribe. A calibration scene is utilized in the main menu to determine user arm length by comparing pose between controllers and HMD. An "Auto Mode" mechanic was enabled to allow users to play at home without research evaluators: the game begins by the user placing their controller on the butterfly for 5 seconds, which loads a set of customized movements that can be remotely set by the therapist or research evaluator. Exercises switch after a default ten repetitions and default 60 second rest period. All variables pertaining to repetitions, exercise type, speed, order, and movement can be customized locally and or remotely through the evaluator interface and prescribed for at-home use or clinical use. This iteration was validated with testing of bi-weekly sessions with five users undergoing upper-extremity rehabilitation over the course of two months [22].

From this work, we found that translating SDT to intrinsically motivate users with the helping behavior mechanic of protecting the butterfly was highly effective in engaging users over the course of two months. Future researchers interested in designing game mechanics may want to consider similar approaches in designing around SDT and helping behavior game mechanics. Additionally our research into emotional intelligence [18, 23] suggests that iVR media can be adapted to existing affective models (e.g. Pleasure-Arousal-Dominance) to increase player engagement by utilizing their in-game behavior. We also found that the motion capture data from off-the-shelf iVR headsets like HTC Vive can be effective for capturing physical rehabilitation success measures (e.g. range of motion and joint torques) when paired with machine learning and biomechanical simulation for greater physical intelligence [16, 2].

1.2 Considerations for Future Work

The COVID-19 pandemic has created an unprecedented challenge for physical rehabilitation due to existing inequities in healthcare and unprepared telehealth technologies. With the past shelter in place orders across the nation, physical rehabilitation clinics had either closed, stopping care for patients altogether, tried to implement social distancing, or adopted telehealth using video conferencing. Yet, while video conferencing enables communication, it often lacks embodied doctor-patient interaction which challenges the doctor to obtain functional health metrics as well as patient motivation and guidance. As the United States is shifting back to in-person rehabilitation amidst the COVID19 Delta Variant in 2021, this problem especially persists for low income and rural communities or highly dense cities that have become “medical deserts” of physical rehabilitation. Consequently, this dissertation research suggests off-the-shelf iVR HMDs such as HTC Vive or Oculus Quest could be incredibly useful for facilitating remote physical rehabilitation from helping users complete exercises to possibly facilitation remote therapist-user interaction. In the future, researchers could extend this dissertation work towards personalizing iVR serious games for other therapy methods, facilitating social interaction (e.g. therapist-patient, patient-patient, and more), adapting more modalities of immersive media for greater emotional engagement (e.g. full-body haptics, olfactory displays), and further training machine learning models to reach more demographics to predict rehabilitation success metrics (e.g. balance oriented physical therapy, lower body exercises). This dissertation research has largely focused on upper-extremity exercise for adults with stroke, developmental disability, and or sports injury. Consequently, it will be critical to include a diverse group of users from all abilities to ensure that future experiences hold greater inclusion, diversity, and equity in assisting both therapists and users with a variety of specialized needs. We believe that iVR systems hold a unique opportunity to empower users and therapists to move beyond the limitations of reality in approaching physical rehabilitation – an opportunity to help make restorative exercise more accessible, affordable, and accurate. While there is undoubtedly more work to be done, this dissertation research points to many potentially fruitful directions for exploring the design and evaluation of iVR experiences for physical rehabilitation. In closing, we have only just scratched the surface of utilizing immersive media for physical care.

Much as Ivan Sutherland described his vision for the ultimate display back in the 1960s, calling researchers to action in bringing it to fruition, we here describe our vision of how the ultimate display could be augmented: The ultimate display for physical rehabilitation would be an immersive virtual reality platform in which the most all human senses are modulated towards exercising. Performing against adversity through challenging yourself to move a once-injured and once-healthy limb would no longer just be a repetitive movement. Such movement can be transformed in the act of protecting a butterfly upon a grassy field; you could feel the raindrops, inhale the smell of the grass, and feel the wind blowing through the bright valley ahead of you.

On some days, the winds would help you move when you could not move any longer. On other days, the winds would follow you through your endeavors into other virtual worlds. Through applying emotional and physical peripherals to an adaptable virtual world for exercise, we can perhaps make the journey of recovery a little more magical.

Bibliography

[1] S. Lessard, P. Pansodtee, A. Robbins, J. M. Trombadore, S. Kurniawan, and M. Teodorescu, “A soft exosuit for flexible upper-extremity rehabilitation,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2018.

[2] 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.

[3] A. Mehrabian, *Basic dimensions for a general psychological theory: Implications for personality, social, environmental, and developmental studies.* Oelgeschlager, Gunn & Hain Cambridge, MA, 1980.

[4] R. F. Bales, T. Adorno, E. Frenkel-Brunswick, D. Levinson, R. Sanford, F. Allport *et al.*, “A comparison between ratings and scorings in the symlog approach.” in *Social Interaction Systems: Theory and Measurement.* Farrar and Rinehart New York, 2001, vol. 34, no. 1, pp. 1–82.

[5] J. A. Russell, “A circumplex model of affect.” *Journal of personality and social psychology*, vol. 39, no. 6, p. 1161, 1980.

[6] A. Mehrabian, *Nonverbal communication.* Routledge, 2017.

[7] M. M. Bradley and P. J. Lang, “Measuring emotion: the self-assessment manikin and the semantic differential,” *Journal of behavior therapy and experimental psychiatry*, vol. 25, no. 1, pp. 49–59, 1994.

[8] P. J. Lang, M. M. Bradley, and B. N. Cuthbert, “International affective picture system (iaps): Technical manual and affective ratings,” *NIMH Center for the Study of Emotion and Attention*, vol. 1, pp. 39–58, 1997.

[9] K. L. Gwet, *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters.* Advanced Analytics, LLC, 2014.

[10] J. Brooke *et al.*, “Sus-a quick and dirty usability scale,” *Usability evaluation in industry*, vol. 189, no. 194, pp. 4–7, 1996.

[11] M. Slater, “Measuring presence: A response to the witmer and singer presence questionnaire,” *Presence*, vol. 8, no. 5, pp. 560–565, 1999.

[12] 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, 2018, pp. 219–225.

[13] 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, 2018.

[14] 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.

[15] 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, 2019, pp. 1448–1456.

[16] A. Elor and S. Kurniawan, “Deep reinforcement learning in immersive virtual reality exergame for agent movement guidance,” in *2020 IEEE 8th International Conference on Serious Games and Applications for Health (SeGAH).* IEEE, 2020, pp. 1–7.

[17] A. Elor and S. Conde, “Exploring the creative possibilities of infinite photogrammetry through spatial computing and extended reality with wave function collapse,” in *In 2020 International Conference on Computational Creativity (ICCC).(Casual Creator Workshop). ACC*, 2020.

[18] 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)(FG)*, pp. 583–587.

[19] 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. 25, 2020.

[20] S. Conde, A. Elor, and S. Kurniawan, “Boundaries: A serious game on relationships for individuals with developmental disabilities,” in *2020 IEEE 8th International Conference on Serious Games and Applications for Health (SeGAH).* IEEE, 2020, pp. 1–7.

[21] T.-E. Vo, R. Jhangiani, A. Robbins, and A. Elor, "Designing user-specific soft robotic wearable muscular interfaces with iterative simulation," in *2020 IEEE International Conference on Smart Computing (SMARTCOMP).* IEEE, 2020, pp. 253–255.

[22] 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.

[23] 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.

[24] 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.

[25] S. Buffington, S. Yu, N. Gersh, and A. Elor, "A stress reactive immersive virtual reality survival game with biofeedback," pp. 1–4, 2021.

[26] R. Jhangiani, A. Elor, M. Kuznetsov, S. Tummala, L. Ott, M. Teodorescu, and S. Kurniawan, "Navigating covid-19 spread in immersive virtual reality: A pilot study on 3d geospatial interaction," pp. 1–4, 2021.

[27] A. Sun, S. Conde, and A. Elor, "Increasing sociability in a virtual world: A serious game for social anxiety disorder," in *2021 IEEE 9th International Conference on Serious Games and Applications for Health (SeGAH).* IEEE, 2021, pp. 1–5.

[28] M. Kuznetsov, A. Elor, S. Kurniawan, C. Bosworth, Y. Rosen, N. Heyer,

M. Teodorescu, B. Paten, and D. Haussler, “The immersive graph genome explorer: Navigating genomics in immersive virtual reality,” in *2021 IEEE 9th International Conference on Serious Games and Applications for Health (SeGAH).* IEEE, 2021, pp. 1–8.

[29] 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.

[30] F. Steinicke, “The science and fiction of the ultimate display,” in *Being really virtual.* Springer, 2016, pp. 19–32.

[31] K. A. Frenkel, “An interview with ivan sutherland,” *Communications of the ACM*, vol. 32, no. 6, pp. 712–714, 1989.

[32] I. E. Sutherland, “The ultimate display,” *Multimedia: From Wagner to virtual reality*, pp. 506–508, 1965.

[33] P. J. Costello, *Health and safety issues associated with virtual reality: a review of current literature.* Advisory Group on Computer Graphics, 1997.

[34] J. D. Westwood, *Medicine Meets Virtual Reality 02/10: Digital Upgrades, Applying Moore’s Law to Health.* IOS Press, 2002, vol. 85.

[35] Statista, “Forecast unit shipments of augmented (ar) and virtual reality (vr) headsets from 2019 to 2023 (in millions),” Statista Research, Tech. Rep., Feb 2020. [Online]. Available: https://www.statista.com/statistics/653390/worldwide-virtual-and-augmented-reality-headset-shipments/

[36] H. Sandler, *Inactivity: physiological effects.* Elsevier, 2012.

[37] L. M. Howden and J. A. Meyer, *Age and sex composition, 2010.* US Department of Commerce, Economics and Statistics Administration, US ..., 2011.

[38] CDC, "Brfss survey data and documentation 2017," C. for Disease Control, Prevention *et al.*, Eds., 2017.

[39] P. Z. Pearce, "Exercise is medicine™," *Current sports medicine reports*, vol. 7, no. 3, pp. 171–175, 2008.

[40] R. Campbell, M. Evans, M. Tucker, B. Quilty, P. Dieppe, and J. Donovan, "Why don't patients do their exercises? understanding non-compliance with physiotherapy in patients with osteoarthritis of the knee," *Journal of epidemiology and community health*, vol. 55, no. 2, pp. 132–138, 2001.

[41] K. Jack, S. M. McLean, J. K. Moffett, and E. Gardiner, "Barriers to treatment adherence in physiotherapy outpatient clinics: a systematic review," *Manual therapy*, vol. 15, no. 3, pp. 220–228, 2010.

[42] 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.

[43] G. C. Burdea, "Virtual rehabilitation–benefits and challenges," *Methods of information in medicine*, vol. 42, no. 05, pp. 519–523, 2003.

[44] E. M. Sluijs, G. J. Kok, and J. Van der Zee, "Correlates of exercise compliance in physical therapy," *Physical therapy*, vol. 73, no. 11, pp. 771–782, 1993.

[45] R. Mellecker and A. McManus, "Active video games and physical activity recommendations: A comparison of the gamercize stepper, xbox kinect and xavix

j-mat," *Journal of Science and Medicine in Sport*, vol. 17, no. 3, pp. 288–292, 2014.

[46] N. N. Byl, G. M. Abrams, E. Pitsch, I. Fedulow, H. Kim, M. Simkins, S. Nagarajan, and J. Rosen, "Chronic stroke survivors achieve comparable outcomes following virtual task specific repetitive training guided by a wearable robotic orthosis (ul-exo7) and actual task specific repetitive training guided by a physical therapist," *Journal of Hand Therapy*, vol. 26, no. 4, pp. 343–352, 2013.

[47] J. E. Deutsch, J. A. Lewis, and G. Burdea, "Technical and patient performance using a virtual reality-integrated telerehabilitation system: preliminary finding," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 15, no. 1, pp. 30–35, 2007.

[48] S. J. M. Bamberg, A. Y. Benbasat, D. M. Scarborough, D. E. Krebs, and J. A. Paradiso, "Gait analysis using a shoe-integrated wireless sensor system," *IEEE transactions on information technology in biomedicine*, vol. 12, no. 4, pp. 413–423, 2008.

[49] R. W. Lindeman, Y. Yanagida, K. Hosaka, and S. Abe, "The tactapack: A wireless sensor/actuator package for physical therapy applications," in *Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2006 14th Symposium on.* IEEE, 2006, pp. 337–341.

[50] 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.

[51] J. J. LaViola Jr, "A discussion of cybersickness in virtual environments," *ACM Sigchi Bulletin*, vol. 32, no. 1, pp. 47–56, 2000.

[52] 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/

[53] M. E. Seligman *et al.*, "Positive psychology, positive prevention, and positive therapy," *Handbook of positive psychology*, vol. 2, no. 2002, pp. 3–12, 2002.

[54] G. A. d. Assis, A. G. D. Corrêa, M. B. R. Martins, W. G. Pedrozo, and R. d. D. Lopes, "An augmented reality system for upper-limb post-stroke motor rehabilitation: a feasibility study," *Disability and Rehabilitation: Assistive Technology*, vol. 11, no. 6, pp. 521–528, 2016.

[55] D. E. Levac, S. M. Glegg, H. Sveistrup, H. Colquhoun, P. Miller, H. Finestone, V. DePaul, J. E. Harris, and D. Velikonja, "Promoting therapists' use of motor learning strategies within virtual reality-based stroke rehabilitation," *PloS one*, vol. 11, no. 12, p. e0168311, 2016.

[56] A. Baldominos, Y. Saez, and C. G. del Pozo, "An approach to physical rehabilitation using state-of-the-art virtual reality and motion tracking technologies," *Procedia Computer Science*, vol. 64, pp. 10–16, 2015.

[57] Y. Salem, S. J. Gropack, D. Coffin, and E. M. Godwin, "Effectiveness of a low-cost virtual reality system for children with developmental delay: a preliminary randomised single-blind controlled trial," *Physiotherapy*, vol. 98, no. 3, pp. 189–195, 2012.

[58] S. Straudi, G. Severini, A. S. Charabati, C. Pavarelli, G. Gamberini, A. Scotti, and N. Basaglia, "The effects of video game therapy on balance and attention in chronic ambulatory traumatic brain injury: an exploratory study," *BMC neurology*, vol. 17, no. 1, p. 86, 2017.

[59] M. R. Kandalaft, N. Didehbani, D. C. Krawczyk, T. T. Allen, and S. B. Chapman, "Virtual reality social cognition training for young adults with high-functioning autism," *Journal of autism and developmental disorders*, vol. 43, no. 1, pp. 34–44, 2013.

[60] D. Freeman, S. Reeve, A. Robinson, A. Ehlers, D. Clark, B. Spanlang, and M. Slater, "Virtual reality in the assessment, understanding, and treatment of mental health disorders," *Psychological medicine*, vol. 47, no. 14, pp. 2393–2400, 2017.

[61] 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.

[62] 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.

[63] 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.

[64] 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.

[65] 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.

[66] 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.

[67] 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.

[68] 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.

[69] 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.

[70] 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.

[71] 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.

[72] 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.

[73] 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.

[74] 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.

[75] 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.

[76] C. Cruz-Neira, D. J. Sandin, and T. A. DeFanti, "Surround-screen projection-

based virtual reality: the design and implementation of the cave," in *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, 1993, pp. 135–142.

[77] 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.

[78] J. Dascal, M. Reid, W. W. IsHak, B. Spiegel, J. Recacho, B. Rosen, and I. Danovitch, "Virtual reality and medical inpatients: A systematic review of randomized, controlled trials," *Innovations in clinical neuroscience*, vol. 14, no. 1-2, p. 14, 2017.

[79] P. S. Lum, G. Uswatte, E. Taub, P. Hardin, and V. W. Mark, "A telerehabilitation approach to delivery of constraint-induced movement therapy," *Journal of rehabilitation research and development*, vol. 43, no. 3, p. 391, 2006.

[80] D. Kairy, M. Tousignant, N. Leclerc, A.-M. Côté, M. Levasseur *et al.*, "The patient's perspective of in-home telerehabilitation physiotherapy services following total knee arthroplasty," *International journal of environmental research and public health*, vol. 10, no. 9, pp. 3998–4011, 2013.

[81] L. Piron, A. Turolla, M. Agostini, C. Zucconi, F. Cortese, M. Zampolini, M. Zannini, M. Dam, L. Ventura, M. Battauz *et al.*, "Exercises for paretic upper limb after stroke: a combined virtual-reality and telemedicine approach," *Journal of Rehabilitation Medicine*, vol. 41, no. 12, pp. 1016–1020, 2009.

[82] R. Lloréns, E. Noé, C. Colomer, and M. Alcañiz, "Effectiveness, usability, and

cost-benefit of a virtual reality–based telerehabilitation program for balance recovery after stroke: A randomized controlled trial," *Archives of physical medicine and rehabilitation*, vol. 96, no. 3, pp. 418–425, 2015.

[83] T. D. Parsons and A. A. Rizzo, "Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis," *Journal of behavior therapy and experimental psychiatry*, vol. 39, no. 3, pp. 250–261, 2008.

[84] J. I. Gold, S. H. Kim, A. J. Kant, M. H. Joseph, and A. S. Rizzo, "Effectiveness of virtual reality for pediatric pain distraction during iv placement," *CyberPsychology & Behavior*, vol. 9, no. 2, pp. 207–212, 2006.

[85] A. Rizzo, A. Hartholt, M. Grimani, A. Leeds, and M. Liewer, "Virtual reality exposure therapy for combat-related posttraumatic stress disorder," *Computer*, vol. 47, no. 7, pp. 31–37, 2014.

[86] M. Slater, "Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 364, no. 1535, pp. 3549–3557, 2009.

[87] R. Skarbez, F. P. Brooks, Jr, and M. C. Whitton, "A survey of presence and related concepts," *ACM Computing Surveys (CSUR)*, vol. 50, no. 6, pp. 1–39, 2017.

[88] C. J. Bohil, B. Alicea, and F. A. Biocca, "Virtual reality in neuroscience research and therapy," *Nature reviews neuroscience*, vol. 12, no. 12, pp. 752–762, 2011.

[89] M. Lombard, T. B. Ditton, D. Crane, B. Davis, G. Gil-Egui, K. Horvath, J. Rossman, and S. Park, "Measuring presence: A literature-based approach to the de-

velopment of a standardized paper-and-pencil instrument," in *Third international workshop on presence, delft, the netherlands*, vol. 240, 2000, pp. 2–4.

[90] D. A. Bowman and R. P. McMahan, "Virtual reality: how much immersion is enough?" *Computer*, vol. 40, no. 7, pp. 36–43, 2007.

[91] 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.

[92] L. Mertz, "Virtual reality is taking the hurt out of pain," *IEEE pulse*, vol. 10, no. 3, pp. 3–8, 2019.

[93] T. Schubert, F. Friedmann, and H. Regenbrecht, "The experience of presence: Factor analytic insights," *Presence: Teleoperators & Virtual Environments*, vol. 10, no. 3, pp. 266–281, 2001.

[94] B. G. Witmer and M. J. Singer, "Measuring presence in virtual environments: A presence questionnaire," *Presence*, vol. 7, no. 3, pp. 225–240, 1998.

[95] A. K. Seth, K. Suzuki, and H. D. Critchley, "An interoceptive predictive coding model of conscious presence," *Frontiers in psychology*, vol. 2, p. 395, 2012.

[96] M. J. Schuemie, P. Van Der Straaten, M. Krijn, and C. A. Van Der Mast, "Research on presence in virtual reality: A survey," *CyberPsychology & Behavior*, vol. 4, no. 2, pp. 183–201, 2001.

[97] P. Salovey, A. J. Rothman, J. B. Detweiler, and W. T. Steward, "Emotional states and physical health." *American psychologist*, vol. 55, no. 1, p. 110, 2000.

[98] L. S. Richman, L. Kubzansky, J. Maselko, I. Kawachi, P. Choo, and M. Bauer, "Positive emotion and health: going beyond the negative." *Health psychology*, vol. 24, no. 4, p. 422, 2005.

[99] P. Ekman, "An argument for basic emotions," *Cognition & emotion*, vol. 6, no. 3-4, pp. 169–200, 1992.

[100] C. Mohn, H. Argstatter, and F.-W. Wilker, "Perception of six basic emotions in music," *Psychology of Music*, vol. 39, no. 4, pp. 503–517, 2011.

[101] C. Collet, E. Vernet-Maury, and A. Dittmar, "Autonomic nervous system response pattern specificity to basic emotions," *International Journal of Psychophysiology*, vol. 1, no. 25, pp. 53–54, 1997.

[102] R. W. Picard, *Affective computing.* MIT press, 2000.

[103] K. Kim, "Emotion modeling and machine learning in affective computing," *Preprint at https://api.semanticscholar.org/CorpusID:239954*, 2014.

[104] B. Meuleman and D. Rudrauf, "Induction and profiling of strong multi-componential emotions in virtual reality," *IEEE Transactions on Affective Computing*, 2018.

[105] Y. Liu, O. Sourina, and M. K. Nguyen, "Real-time eeg-based emotion recognition and its applications," in *Transactions on computational science XII.* Springer, 2011, pp. 256–277.

[106] T. Waltemate, D. Gall, D. Roth, M. Botsch, and M. E. Latoschik, "The impact of avatar personalization and immersion on virtual body ownership, presence, and

emotional response," *IEEE transactions on visualization and computer graphics*, vol. 24, no. 4, pp. 1643–1652, 2018.

[107] B. Geethanjali, K. Adalarasu, A. Hemapraba, S. Pravin Kumar, and R. Rajasekeran, "Emotion analysis using sam (self-assessment manikin) scale." *Biomedical Research (0970-938X)*, vol. 28, 2017.

[108] T.-M. Bynion and M. T. Feldner, "Self-assessment manikin," *Encyclopedia of personality and individual differences*, pp. 1–3, 2017.

[109] N. Yee and J. Bailenson, "The proteus effect: The effect of transformed self-representation on behavior," *Human communication research*, vol. 33, no. 3, pp. 271–290, 2007.

[110] F. A. Geldard, R. O'Hehir, and D. Gavens, *The human senses.* Wiley New York, 1953.

[111] J. N. Bailenson, N. Yee, S. Brave, D. Merget, and D. Koslow, "Virtual interpersonal touch: expressing and recognizing emotions through haptic devices," *Human–Computer Interaction*, vol. 22, no. 3, pp. 325–353, 2007.

[112] A. Mazzoni and N. Bryan-Kinns, "How does it feel like? an exploratory study of a prototype system to convey emotion through haptic wearable devices," in *2015 7th International Conference on Intelligent Technologies for Interactive Entertainment (INTETAIN).* IEEE, 2015, pp. 64–68.

[113] D. Bonnet, M. Ammi, and J.-C. Martin, "Improvement of the recognition of facial expressions with haptic feedback," in *2011 IEEE International Workshop on Haptic Audio Visual Environments and Games.* IEEE, 2011, pp. 81–87.

[114] K. Salminen, V. Surakka, J. Lylykangas, J. Raisamo, R. Saarinen, R. Raisamo, J. Rantala, and G. Evreinov, "Emotional and behavioral responses to haptic stimulation," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2008, pp. 1555–1562.

[115] M. Obrist, S. Subramanian, E. Gatti, B. Long, and T. Carter, "Emotions mediated through mid-air haptics," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2015, pp. 2053–2062.

[116] P. Miri, R. Flory, A. Uusberg, H. Culbertson, R. H. Harvey, A. Kelman, D. E. Peper, J. J. Gross, K. Isbister, and K. Marzullo, "Piv: Placement, pattern, and personalization of an inconspicuous vibrotactile breathing pacer," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 27, no. 1, pp. 1–44, 2020.

[117] K. Fox, "The smell report," *Social Issues Research Centre*, 2009.

[118] A. Hirsch and J. Gruss, "Human male sexual response to olfactory stimuli," *J Neurol Orthop Med Surg*, vol. 19, pp. 14–19, 1999.

[119] F. Biocca, J. Kim, and Y. Choi, "Visual touch in virtual environments: An exploratory study of presence, multimodal interfaces, and cross-modal sensory illusions," *Presence: Teleoperators & Virtual Environments*, vol. 10, no. 3, pp. 247–265, 2001.

[120] T. Bernard, A. Gonzalez, V. Miale, K. Vangara, L. Stephane, and W. E. Scott, "Haptic feedback astronaut suit for mitigating extra-vehicular activity spatial disorientation," in *AIAA SPACE and Astronautics Forum and Exposition*, 2017, p. 5113.

[121] L. Goedschalk, T. Bosse, and M. Otte, "Get your virtual hands off me!–developing threatening ivas using haptic feedback," in *Benelux Conference on Artificial Intelligence.* Springer, 2017, pp. 61–75.

[122] C. Krogmeier, C. Mousas, and D. Whittinghill, "Human, virtual human, bump! a preliminary study on haptic feedback," in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR).* IEEE, 2019, pp. 1032–1033.

[123] D. Wolf, M. Rietzler, L. Hnatek, and E. Rukzio, "Face/on: Multi-modal haptic feedback for head-mounted displays in virtual reality," *IEEE transactions on visualization and computer graphics*, vol. 25, no. 11, pp. 3169–3177, 2019.

[124] R. L. Peiris, W. Peng, Z. Chen, L. Chan, and K. Minamizawa, "Thermovr: Exploring integrated thermal haptic feedback with head mounted displays," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems.* ACM, 2017, pp. 5452–5456.

[125] E. Doukakis, K. Debattista, T. Bashford-Rogers, A. Dhokia, A. Asadipour, A. Chalmers, and C. Harvey, "Audio-visual-olfactory resource allocation for tri-modal virtual environments," *IEEE transactions on visualization and computer graphics*, vol. 25, no. 5, pp. 1865–1875, 2019.

[126] D. Warnock, M. McGee-Lennon, and S. Brewster, "The role of modality in notification performance," in *IFIP Conference on Human-Computer Interaction.* Springer, 2011, pp. 572–588.

[127] M. R. McGee-Lennon and S. Brewster, "Reminders that make sense: Designing multimodal notifications for the home," in *2011 5th International Conference on*

Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops. IEEE, 2011, pp. 495–501.

[128] M. Azmandian, M. Hancock, H. Benko, E. Ofek, and A. D. Wilson, "Haptic retargeting: Dynamic repurposing of passive haptics for enhanced virtual reality experiences," in *Proceedings of the 2016 chi conference on human factors in computing systems.* ACM, 2016, pp. 1968–1979.

[129] M. Ischer, N. Baron, C. Mermoud, I. Cayeux, C. Porcherot, D. Sander, and S. Delplanque, "How incorporation of scents could enhance immersive virtual experiences," *Frontiers in psychology*, vol. 5, p. 736, 2014.

[130] M. P. Aiken and M. J. Berry, "Posttraumatic stress disorder: possibilities for olfaction and virtual reality exposure therapy," *Virtual Reality*, vol. 19, no. 2, pp. 95–109, 2015.

[131] T. Schweizer, F. Renner, D. Sun, B. Kleim, E. A. Holmes, and B. Tuschen-Caffier, "Psychophysiological reactivity, coping behaviour and intrusive memories upon multisensory virtual reality and script-driven imagery analogue trauma: A randomised controlled crossover study," *Journal of anxiety disorders*, vol. 59, pp. 42–52, 2018.

[132] H. Q. Dinh, N. Walker, L. F. Hodges, C. Song, and A. Kobayashi, "Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments," in *Proceedings IEEE Virtual Reality (Cat. No. 99CB36316).* IEEE, 1999, pp. 222–228.

[133] R. Soares, E. Siqueira, M. Miura, T. Silva, and C. Castanho, "Biofeedback sen-

sors in game telemetry research," *Simpósio Brasileiro de Jogos e Entretenimento Digital*, pp. 81–89, 2016.

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

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

[136] 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.

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

[138] 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.

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

[140] 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.

[141] 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.

[142] 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.

[143] 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.

[144] 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.

[145] 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.

[146] 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.

[147] G. Deuschl, A. Eisen *et al.*, "Recommendations for the practice of clinical neurophysiology: guidelines of the international federation of clinical neurophysiology," 1999.

[148] 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.

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

[150] 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.

[151] 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.

[152] 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.

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

[154] 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.

[155] J. A. Hobson and E. F. Pace-Schott, "The cognitive neuroscience of sleep: neuronal systems, consciousness and learning," *Nature Reviews Neuroscience*, vol. 3, no. 9, p. 679, 2002.

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

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

[158] 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.

[159] 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.

[160] 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.

[161] 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.

[162] 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.

[163] 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.

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

[165] 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.

[166] 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.

[167] M. Eimer, A. Holmes, and F. P. McGlone, "The role of spatial attention in the processing of facial expression: an erp study of rapid brain responses to six basic emotions," *Cognitive, Affective, & Behavioral Neuroscience*, vol. 3, no. 2, pp. 97–110, 2003.

[168] S. B. i Badia, L. V. Quintero, M. S. Cameirao, A. Chirico, S. Triberti, P. Cipresso, and A. Gaggioli, "Towards emotionally-adaptive virtual reality for mental health applications," *IEEE journal of biomedical and health informatics*, 2018.

[169] W. H. Redd, S. L. Manne, B. Peters, P. B. Jacobsen, and H. Schmidt, "Fragrance administration to reduce anxiety during mr imaging," *Journal of Magnetic Resonance Imaging*, vol. 4, no. 4, pp. 623–626, 1994.

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

[171] J. Amores, R. Richer, N. Zhao, P. Maes, and B. M. Eskofier, "Promoting relaxation using virtual reality, olfactory interfaces and wearable eeg," in *2018 IEEE 15th international conference on wearable and implantable body sensor networks (BSN).* IEEE, 2018, pp. 98–101.

[172] H. B. Abdessalem, M. Boukadida, and C. Frasson, "Virtual reality game adaptation using neurofeedback," in *The Thirty-First International Flairs Conference*, 2018.

[173] J. Marín-Morales, J. L. Higuera-Trujillo, A. Greco, J. Guixeres, C. Llinares, E. P. Scilingo, M. Alcañiz, and G. Valenza, "Affective computing in virtual reality: emotion recognition from brain and heartbeat dynamics using wearable sensors," *Scientific reports*, vol. 8, no. 1, p. 13657, 2018.

[174] D. Krönert, A. Grünewald, F. Li, M. Grzegorzek, and R. Brück, "Sensor headband for emotion recognition in a virtual reality environment," in *International Conference on Information Technologies in Biomedicine.* Springer, 2018, pp. 539–548.

[175] M. Van Rooij, A. Lobel, O. Harris, N. Smit, and I. Granic, "Deep: A biofeedback virtual reality game for children at-risk for anxiety," in *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems.* ACM, 2016, pp. 1989–1997.

[176] A. Chalmers and A. Ferko, "Levels of realism: From virtual reality to real virtuality," in *Proceedings of the 24th Spring Conference on Computer Graphics.* ACM, 2008, pp. 19–25.

[177] K. H. Teigen, “Yerkes-dodson: A law for all seasons,” *Theory & Psychology*, vol. 4, no. 4, pp. 525–547, 1994.

[178] M. Csikszentmihalyi, S. Abuhamdeh, J. Nakamura *et al.*, “Flow,” 1990.

[179] M. Csikszentmihalyi, “Beyond boredom and anxiety: The experience of play in work and leisure,” 1975.

[180] P. Sweetser and P. Wyeth, “Gameflow: a model for evaluating player enjoyment in games,” *Computers in Entertainment (CIE)*, vol. 3, no. 3, pp. 3–3, 2005.

[181] S. H.-W. Chuah, “Why and who will adopt extended reality technology? literature review, synthesis, and future research agenda,” *Literature Review, Synthesis, and Future Research Agenda (December 13, 2018)*, 2018.

[182] M.-S. Bracq, E. Michinov, and P. Jannin, “Virtual reality simulation in nontechnical skills training for healthcare professionals: A systematic review,” *Simulation in Healthcare*, vol. 14, no. 3, pp. 188–194, 2019.

[183] R. Yung and C. Khoo-Lattimore, “New realities: a systematic literature review on virtual reality and augmented reality in tourism research,” *Current Issues in Tourism*, vol. 22, no. 17, pp. 2056–2081, 2019.

[184] T. Rose, C. S. Nam, and K. B. Chen, “Immersion of virtual reality for rehabilitation-review,” *Applied ergonomics*, vol. 69, pp. 153–161, 2018.

[185] L. P. Berg and J. M. Vance, “Industry use of virtual reality in product design and manufacturing: a survey,” *Virtual reality*, vol. 21, no. 1, pp. 1–17, 2017.

[186] Å. Fast-Berglund, L. Gong, and D. Li, “Testing and validating extended reality

(xr) technologies in manufacturing," *Procedia Manufacturing*, vol. 25, pp. 31–38, 2018.

[187] F. Bonetti, G. Warnaby, and L. Quinn, "Augmented reality and virtual reality in physical and online retailing: A review, synthesis and research agenda," in *Augmented reality and virtual reality.* Springer, 2018, pp. 119–132.

[188] P. T. Jaeger, "Disability, human rights, and social justice: The ongoing struggle for online accessibility and equality," *First Monday*, 2015.

[189] A. Board, "Laws concerning the access board," *United States Access Board - Advancing Full Access and Inclusion for All. Internet: https://www.access-board.gov/the-board/laws [Accessed on June. 27, 2020]*, 2020.

[190] G. M. Greco, "On accessibility as a human right, with an application to media accessibility," in *Researching audio description.* Springer, 2016, pp. 11–33.

[191] R. F. Erlandson, *Universal and accessible design for products, services, and processes.* CRC Press, 2007.

[192] L. H. Pelled, "Demographic diversity, conflict, and work group outcomes: An intervening process theory," *Organization science*, vol. 7, no. 6, pp. 615–631, 1996.

[193] XRA, "Xr access: A community committed to making virtual, augmented, and mixed reality (xr) accessible to people with disabilities," *XR Access. Internet: https://xraccess.org/about/ [Accessed on June. 25, 2020]*, feb 2020.

[194] Y. Zhao, E. Cutrell, C. Holz, M. R. Morris, E. Ofek, and A. D. Wilson, "Seeingvr: A set of tools to make virtual reality more accessible to people with low vision,"

in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–14.

[195] M. J. Habgood, D. Moore, D. Wilson, and S. Alapont, "Rapid, continuous movement between nodes as an accessible virtual reality locomotion technique," in *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR).* IEEE, 2018, pp. 371–378.

[196] A. Mower, R. Nguyen, and K. Frank, "Evaluation of technology accessibility and user sentiment in learning through virtual reality modality," in *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020, pp. 1–8.

[197] F. Garzotto, V. Matarazzo, N. Messina, M. Gelsomini, and C. Riva, "Improving museum accessibility through storytelling in wearable immersive virtual reality," in *2018 3rd Digital Heritage International Congress (DigitalHERITAGE) held jointly with 2018 24th International Conference on Virtual Systems & Multimedia (VSMM 2018).* IEEE, 2018, pp. 1–8.

[198] Z. Rashid, J. Melià-Seguí, R. Pous, and E. Peig, "Using augmented reality and internet of things to improve accessibility of people with motor disabilities in the context of smart cities," *Future Generation Computer Systems*, vol. 76, pp. 248–261, 2017.

[199] J. M. Coughlan and J. Miele, "Ar4vi: Ar as an accessibility tool for people with visual impairments," in *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct).* IEEE, 2017, pp. 288–292.

[200] S. M. S. Ferdous, "Improve accessibility of virtual and augmented reality for people with balance impairments," in *2017 IEEE Virtual Reality (VR)*. IEEE, 2017, pp. 421–422.

[201] Y.-J. Chang, Y.-S. Kang, and P.-C. Huang, "An augmented reality (ar)-based vocational task prompting system for people with cognitive impairments," *Research in developmental disabilities*, vol. 34, no. 10, pp. 3049–3056, 2013.

[202] D. D. McMahon, D. F. Cihak, R. E. Wright, and S. M. Bell, "Augmented reality for teaching science vocabulary to postsecondary education students with intellectual disabilities and autism," *Journal of Research on Technology in Education*, vol. 48, no. 1, pp. 38–56, 2016.

[203] N. Fallah, "Audionav: a mixed reality navigation system for individuals who are visually impaired," *ACM SIGACCESS Accessibility and Computing*, no. 96, pp. 24–27, 2010.

[204] P. O. Hedvall, "Towards the era of mixed reality: accessibility meets three waves of hci," in *Symposium of the Austrian HCI and Usability Engineering Group*. Springer, 2009, pp. 264–278.

[205] D. Schneider, A. Otte, T. Gesslein, P. Gagel, B. Kuth, M. S. Damlakhi, O. Dietz, E. Ofek, M. Pahud, P. O. Kristensson *et al.*, "Reconviguration: Reconfiguring physical keyboards in virtual reality," *IEEE transactions on visualization and computer graphics*, vol. 25, no. 11, pp. 3190–3201, 2019.

[206] A. F. Siu, M. Sinclair, R. Kovacs, E. Ofek, C. Holz, and E. Cutrell, "Virtual reality without vision: A haptic and auditory white cane to navigate complex

virtual worlds," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020, pp. 1–13.

[207] E. Strasnick, C. Holz, E. Ofek, M. Sinclair, and H. Benko, "Haptic links: Bimanual haptics for virtual reality using variable stiffness actuation," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018, pp. 1–12.

[208] J. Yang, C. Holz, E. Ofek, and A. D. Wilson, "Dreamwalker: Substituting real-world walking experiences with a virtual reality," in *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, 2019, pp. 1093–1107.

[209] P. De Groote, "Accessibility features on bnp paribas fortis atms," in *International Conference on Computers Helping People with Special Needs.* Springer, 2018, pp. 303–306.

[210] N. L. Shaheen and S. Lohnes Watulak, "Bringing disability into the discussion: Examining technology accessibility as an equity concern in the field of instructional technology," *Journal of Research on Technology in Education*, vol. 51, no. 2, pp. 187–201, 2019.

[211] A. Guo, J. Kong, M. Rivera, F. F. Xu, and J. P. Bigham, "Statelens: A reverse engineering solution for making existing dynamic touchscreens accessible," in *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, 2019, pp. 371–385.

[212] E. Steinfeld, J. Maisel, D. Feathers, and C. D'Souza, "Anthropometry and stan-

dards for wheeled mobility: an international comparison," *Assistive Technology®*, vol. 22, no. 1, pp. 51–67, 2010.

[213] L. Chelkowski, Z. Yan, and K. Asaro-Saddler, "The use of mobile devices with students with disabilities: a literature review," *Preventing School Failure: Alternative Education for Children and Youth*, vol. 63, no. 3, pp. 277–295, 2019.

[214] L. S. Sullivan, E. Klein, T. Brown, M. Sample, M. Pham, P. Tubig, R. Folland, A. Truitt, and S. Goering, "Keeping disability in mind: a case study in implantable brain–computer interface research," *Science and engineering ethics*, vol. 24, no. 2, pp. 479–504, 2018.

[215] C. Guger, B. Z. Allison, and N. Mrachacz-Kersting, "Brain-computer interface research: A state-of-the-art summary 7," in *Brain-Computer Interface Research*. Springer, 2019, pp. 1–9.

[216] R. A. Ramadan and A. V. Vasilakos, "Brain computer interface: control signals review," *Neurocomputing*, vol. 223, pp. 26–44, 2017.

[217] P. Cairns, C. Power, M. Barlet, and G. Haynes, "Future design of accessibility in games: A design vocabulary," *International Journal of Human-Computer Studies*, vol. 131, pp. 64–71, 2019.

[218] W3C, "Inclusive design for immersive web standards: W3c workshop report," *World Wide Web Consortium (W3C). Internet: https://www.w3.org/2019/08/inclusive-xr-workshop/report.html [Accessed on June. 27, 2020]*, November 2019.

[219] ——, "W3c workshop on web virtual reality: Workshop report," *World Wide*

Web Consortium (W3C). Internet: https://www.w3.org/2016/06/vr-workshop/report.html [Accessed on June. 27, 2020], October 2016.

[220] ——, "Webxr standards and accessibility architecture issues," *World Wide Web Consortium (W3C). Internet: https://www.w3.org/WAI/APA/wiki/WebXR_Standards_and_Accessibility_Architecture_Issues [Accessed on June. 27, 2020]*, February 2020.

[221] ——, "Xaur - user needs and requirements for augmented and virtual reality [draft]," *World Wide Web Consortium (W3C). Internet: https://www.w3.org/WAI/APA/wiki/Xaur_draft [Accessed on June. 27, 2020]*, December 2019.

[222] X. A. (CRA), "Xra: The xra's mission is to promote responsible development and thoughtful advancement of xr that foster positive societal outcomes." *Internet: https://xra.org/ [Accessed on June. 27, 2020]*, June 2020.

[223] K. Group, "Unifying reality: Openxr is a royalty-free, open standard that provides high-performance access to augmented reality (ar) and virtual reality (vr)—collectively known as xr—platforms and devices." *Internet: https://www.khronos.org/openxr/ [Accessed on June. 27, 2020]*, December 2016.

[224] M. Mott, E. Cutrell, M. G. Franco, C. Holz, E. Ofek, R. Stoakley, and M. R. Morris, "Accessible by design: An opportunity for virtual reality," in *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, 2019, pp. 451–454.

[225] K. Nakamura, "My algorithms have determined you're not human: Ai-ml, re-

verse turing-tests, and the disability experience," in *The 21st International ACM SIGACCESS Conference on Computers and Accessibility*, 2019, pp. 1–2.

[226] S. Azenkot, L. Goldberg, J. Taft, and S. Soloway, "Xr symposium report," *XR Access. Internet: https://xraccess.org/category/reports/ [Accessed on June. 25, 2020]*, 2019.

[227] 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.*

[228] M. Ora 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.

[229] J. Miller-Merrell, "Podcast: How xr is changing workplace training and accessibility. featuring joel ward." *Partnership on Employment & Accessible Technology (PEAT). Internet: https://www.peatworks.org/content/podcast-how-xr-changing-workplace-training-and-accessibility [Accessed on June. 25, 2020]*, mar 2020.

[230] J. Ward, "Let's make xr accessible," *Joelsef Explains It All. Internet: https://joelsef.com/lets-make-xr-accessible [Accessed on June. 25, 2020]*, jan 2020.

[231] S. Armstrong, "Adaptive controller is getting people with disabilities back into gaming," *Wired. Retrieved from www.wired.co.uk/article/microsoft-xbox-adaptive-controller*, 2018.

[232] M. Dombrowski, P. A. Smith, A. Manero, and J. Sparkman, “Designing inclusive virtual reality experiences,” in *International Conference on Human-Computer Interaction.* Springer, 2019, pp. 33–43.

[233] M. Bianchin and A. Heylighen, “Fair by design. addressing the paradox of inclusive design approaches,” *The Design Journal*, vol. 20, no. sup1, pp. S3162–S3170, 2017.

[234] 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, 2019, pp. 1448–1456.

[235] S. K. Carter and J. A. Rizzo, “Use of outpatient physical therapy services by people with musculoskeletal conditions,” *Physical therapy*, vol. 87, no. 5, p. 497, 2007.

[236] “U.s. department of health and human services centers for medicare & medicaid services (cms). report to the congress: Medicare payment policy. technical report, 2016.” March 2016, retrieved February 7, 2016 from http://www.medpac.gov/docs/default-source/reports/march-2016-report-to-the-congress-medicare-payment-policy.pdf?sfvrsn=0.

[237] D. K. Chan, C. Lonsdale, P. Y. Ho, P. S. Yung, and K. M. Chan, “Patient motivation and adherence to postsurgery rehabilitation exercise recommendations: the influence of physiotherapists’ autonomy-supportive behaviors,” *Archives of physical medicine and rehabilitation*, vol. 90, no. 12, pp. 1977–1982, 2009.

[238] A. K. Roy, Y. Soni, and S. Dubey, “Enhancing effectiveness of motor rehabilita-

tion using kinect motion sensing technology," in *Global Humanitarian Technology Conference: South Asia Satellite (GHTC-SAS), 2013 IEEE.* IEEE, 2013, pp. 298–304.

[239] V. W. Mark and E. Taub, "Constraint-induced movement therapy for chronic stroke hemiparesis and other disabilities," *Restorative neurology and neuroscience*, vol. 22, no. 3-5, pp. 317–336, 2004.

[240] E.-K. Ji and S.-H. Lee, "Effects of virtual reality training with modified constraint-induced movement therapy on upper extremity function in acute stage stroke: a preliminary study," *Journal of physical therapy science*, vol. 28, no. 11, pp. 3168–3172, 2016.

[241] M. da Silva Cameirão, S. Bermúdez i Badia, E. Duarte, and P. F. Verschure, "Virtual reality based rehabilitation speeds up functional recovery of the upper extremities after stroke: a randomized controlled pilot study in the acute phase of stroke using the rehabilitation gaming system," *Restorative neurology and neuroscience*, vol. 29, no. 5, pp. 287–298, 2011.

[242] M. F. Levin, O. Snir, D. G. Liebermann, H. Weingarden, and P. L. Weiss, "Virtual reality versus conventional treatment of reaching ability in chronic stroke: clinical feasibility study," *Neurology and therapy*, vol. 1, no. 1, p. 3, 2012.

[243] A. N. V. Follow, "Virtual reality (vr) continuum - amp new ventures," Dec 2016, https://www.slideshare.net/ampnewventures/virtual-reality-vr-continuum-amp-new-ventures.

[244] G. C. Corp, "Gsv capital corp. the pioneer building," May 2013. [Online]. Available: http://gsvcap.com/wp/market-commentary/project-runway/

[245] S. F. et al., "Feasibility of using the sony playstation 2 gaming platform for an individual poststroke: A case report," *Neurologic Physical Therapy*, vol. 31, pp. 180–189, 2008.

[246] G. Saposnik, R. Teasell, M. Mamdani, J. Hall, W. McIlroy, D. Cheung, K. E. Thorpe, L. G. Cohen, M. Bayley *et al.*, "Effectiveness of virtual reality using wii gaming technology in stroke rehabilitation," *Stroke*, vol. 41, no. 7, pp. 1477–1484, 2010.

[247] S. Heins, S. Dehem, V. Montedoro, F. Rocca, P.-H. de Deken, M. Edwards, B. Dehez, M. Mancas, G. Stoquart, T. Lejeune *et al.*, "Robotic-assisted serious game for motor and cognitive post-stroke rehabilitation," in *the 5th IEEE Conference on Serious Games and Applications for Health*, 2017.

[248] D. Tsoupikova, K. Triandafilou, S. Solanki, A. Barry, F. Preuss, and D. Kamper, "Real-time diagnostic data in multi-user virtual reality post-stroke therapy," in *SIGGRAPH ASIA 2016 VR Showcase*, ser. SA '16. New York, NY, USA: ACM, 2016, pp. 8:1–8:2. [Online]. Available: http://doi.acm.org/10.1145/2996376.2996387

[249] P. Wang, G. C. H. Koh, C. G. Boucharenc, T. M. Xu, Hamasaki, and C. C. Yen, "Developing a tangible gaming board for post-stroke upper limb functional training," in *Proceedings of the Eleventh International Conference on Tangible, Embedded, and Embodied Interaction*, ser. TEI

’17. New York, NY, USA: ACM, 2017, pp. 617–624. [Online]. Available: http://doi.acm.org.oca.ucsc.edu/10.1145/3024969.3025080

[250] M. Pasch, N. Berthouze, B. Dijk, and A. Nijholt, “Motivations, strategies, and movement patterns of video gamers playing nintendo wii boxing,” 2008.

[251] S. J. Page, P. Levine, S. Sisto, Q. Bond, and M. V. Johnston, “Stroke patients’ and therapists’ opinions of constraint-induced movement therapy,” *Clinical rehabilitation*, vol. 16, no. 1, pp. 55–60, 2002.

[252] Y. X. Shi, J. H. Tian, K. H. Yang, and Y. Zhao, “Modified constraint-induced movement therapy versus traditional rehabilitation in patients with upper-extremity dysfunction after stroke: a systematic review and meta-analysis,” *Archives of physical medicine and rehabilitation*, vol. 92, no. 6, pp. 972–982, 2011.

[253] D. T. Staff, “Oculus rift vs. htc vive,” Sep 2017. [Online]. Available: https://www.digitaltrends.com/virtual-reality/oculus-rift-vs-htc-vive/

[254] NINDS, “Post-stroke rehabilitation. nih publication no. 14 1846,” September 2014, retrieved April 30, 2017 from https://stroke.nih.gov/materials/rehabilitation.htm.

[255] O. Kreylos, “Lighthouse tracking examined,” *A developer’s perspective on immersive 3D computer graphics*, May 2016, http://doc-ok.org/?p=1478.

[256] G. Marshall, “Best vr controller: Htc vive vs oculus rift vs playstation vr vs gear vr,” March 2016, retrieved April 30, 2017 from http://www.techradar.com/news/gaming/best-vr-controller-htc-vive-vs-oculus-rift-vs-.

[257] HTC-Corporation, “Vive vr system,” *Vive*, November 2018, https://www.vive.com/us/product/vive-virtual-reality-system/.

[258] F. Soffel, M. Zank, and A. Kunz, “Postural stability analysis in virtual reality using the htc vive,” in *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology.* ACM, 2016, pp. 351–352.

[259] D. C. Niehorster, L. Li, and M. Lappe, “The accuracy and precision of position and orientation tracking in the htc vive virtual reality system for scientific research,” *i-Perception*, vol. 8, no. 3, p. 2041669517708205, 2017.

[260] L. M. Howden and J. A. Meyer, “Age and sex composition: 2010,” *2010 Census Briefs, US Department of Commerce, Economics and Statistics Administration. US CENSUS BUREAU*, 2010.

[261] N. S. Ward, “Compensatory mechanisms in the aging motor system,” *Ageing research reviews*, vol. 5, no. 3, pp. 239–254, 2006.

[262] R. D. Seidler, J. A. Bernard, T. B. Burutolu, B. W. Fling, M. T. Gordon, J. T. Gwin, Y. Kwak, and D. B. Lipps, “Motor control and aging: links to age-related brain structural, functional, and biochemical effects,” *Neuroscience & Biobehavioral Reviews*, vol. 34, no. 5, pp. 721–733, 2010.

[263] R. D. Seidler-Dobrin, J. He, and G. E. Stelmach, “Coactivation to reduce variability in the elderly,” *Motor control*, vol. 2, no. 4, pp. 314–330, 1998.

[264] R. D. Seidler, J. L. Alberts, and G. E. Stelmach, “Changes in multi-joint performance with age,” *Motor control*, vol. 6, no. 1, pp. 19–31, 2002.

[265] N. Shkuratova, M. E. Morris, and F. Huxham, “Effects of age on balance control during walking,” *Archives of physical medicine and rehabilitation*, vol. 85, no. 4, pp. 582–588, 2004.

[266] R. Koopman and L. J. van Loon, "Aging, exercise, and muscle protein metabolism," *Journal of Applied Physiology*, vol. 106, no. 6, pp. 2040–2048, 2009.

[267] "Virtual reality for consumer markets - head-mounted displays, accessory devices, and consumer virtual reality content: Global market analysis and forecasts," Tractica Research, report, Jan 2016. [Online]. Available: https://www.tractica.com/research/virtual-reality-for-consumer-markets/

[268] F. J. Deconinck, A. R. Smorenburg, A. Benham, A. Ledebt, M. G. Feltham, and G. J. Savelsbergh, "Reflections on mirror therapy: a systematic review of the effect of mirror visual feedback on the brain," *Neurorehabilitation and Neural Repair*, vol. 29, no. 4, pp. 349–361, 2015.

[269] M. E. Michielsen, M. Smits, G. M. Ribbers, H. J. Stam, J. N. Van Der Geest, J. B. Bussmann, and R. W. Selles, "The neuronal correlates of mirror therapy: an fmri study on mirror induced visual illusions in patients with stroke," *Journal of neurology, neurosurgery & psychiatry*, vol. 82, no. 4, pp. 393–398, 2011.

[270] V. S. Ramachandran and D. Rogers-Ramachandran, "Synaesthesia in phantom limbs induced with mirrors," *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 263, no. 1369, pp. 377–386, 1996.

[271] H. Thieme, J. Mehrholz, M. Pohl, J. Behrens, and C. Dohle, "Mirror therapy for improving motor function after stroke," *Stroke*, vol. 44, no. 1, pp. e1–e2, 2013.

[272] A. Frisoli, F. Salsedo, M. Bergamasco, B. Rossi, and M. C. Carboncini, "A force-feedback exoskeleton for upper-limb rehabilitation in virtual reality," *Applied Bionics and Biomechanics*, vol. 6, no. 2, pp. 115–126, 2009.

[273] D. Jack, R. Boian, A. S. Merians, M. Tremaine, G. C. Burdea, S. V. Adamovich, M. Recce, and H. Poizner, "Virtual reality-enhanced stroke rehabilitation," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 9, no. 3, pp. 308–318, 2001.

[274] S. J. Housman, V. Le, T. Rahman, R. J. Sanchez, and D. J. Reinkensmeyer, "Arm-training with t-wrex after chronic stroke: preliminary results of a randomized controlled trial," in *Rehabilitation Robotics, 2007. ICORR 2007. IEEE 10th International Conference on.* IEEE, 2007, pp. 562–568.

[275] H. Wang and T. Cochrane, "Mobility impairment, muscle imbalance, muscle weakness, scapular asymmetry and shoulder injury in elite volleyball athletes," *Journal of sports medicine and physical fitness*, vol. 41, no. 3, p. 403, 2001.

[276] S. Lessard, P. Pansodtee, A. Robbins, L. B. Baltaxe-Admony, J. M. Trombadore, M. Teodorescu, A. Agogino, and S. Kurniawan, "Crux: A compliant robotic upper-extremity exosuit for lightweight, portable, multi-joint muscular augmentation," in *Rehabilitation Robotics (ICORR), 2017 International Conference on.* IEEE, 2017, pp. 1633–1638.

[277] A. Hanaor, "Engineering properties of double-layer tensegrity grids," *SpatialStructures at the Tum oftheMillenium*, pp. 195–200, 1991.

[278] M. Wehner, B. Quinlivan, P. M. Aubin, E. Martinez-Villalpando, M. Baumann, L. Stirling, K. Holt, R. Wood, and C. Walsh, "A lightweight soft exosuit for gait assistance," in *Robotics and Automation (ICRA), 2013 IEEE International Conference on.* IEEE, 2013, pp. 3362–3369.

[279] A. Alvara, "Development of an exosuit for rehabilitative use and upper arm force amplification," in *2017 AAAS Annual Meeting (February 16-20, 2017)*. aaas, 2017.

[280] N. Ahmad, R. A. R. Ghazilla, N. M. Khairi, and V. Kasi, "Reviews on various inertial measurement unit (imu) sensor applications," *International Journal of Signal Processing Systems*, vol. 1, no. 2, pp. 256–262, 2013.

[281] ValveSoftware, "Valvesoftware/openvr," *GitHub*, Jul, https://github.com/ValveSoftware/openvr.

[282] F. García-Muro, Á. L. Rodríguez-Fernández, and A. Herrero-de Lucas, "Treatment of myofascial pain in the shoulder with kinesio taping. a case report," *Manual therapy*, vol. 15, no. 3, pp. 292–295, 2010.

[283] D. M. Morris, G. Uswatte, J. E. Crago, E. W. Cook III, and E. Taub, "The reliability of the wolf motor function test for assessing upper extremity function after stroke," *Archives of physical medicine and rehabilitation*, vol. 82, no. 6, pp. 750–755, 2001.

[284] D. J. Gladstone, C. J. Danells, and S. E. Black, "The fugl-meyer assessment of motor recovery after stroke: a critical review of its measurement properties," *Neurorehabilitation and neural repair*, vol. 16, no. 3, pp. 232–240, 2002.

[285] A. Mehrabian, "Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament," *Current Psychology*, vol. 14, no. 4, pp. 261–292, 1996.

[286] M. F. Floyd, “Pleasure, arousal, and dominance: Exploring affective determinants of recreation satisfaction,” *Leisure Sciences*, vol. 19, no. 2, pp. 83–96, 1997.

[287] J. D. Morris, “Observations: Sam: the self-assessment manikin; an efficient cross-cultural measurement of emotional response,” *Journal of advertising research*, vol. 35, no. 6, pp. 63–68, 1995.

[288] P. J. Lang, M. M. Bradley, and B. N. Cuthbert, “International affective picture system (iaps): Technical manual and affective ratings,” *NIMH Center for the Study of Emotion and Attention*, vol. 1, pp. 39–58, 1997.

[289] P. Lang and M. M. Bradley, “The international affective picture system (iaps) in the study of emotion and attention,” *Handbook of emotion elicitation and assessment*, vol. 29, 2007.

[290] P. J. Standen and D. J. Brown, “Virtual reality in the rehabilitation of people with intellectual disabilities,” *Cyberpsychology & behavior*, vol. 8, no. 3, pp. 272–282, 2005.

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

[292] I. Bakker, T. van der Voordt, P. Vink, and J. de Boon, “Pleasure, arousal, dominance: Mehrabian and russell revisited,” *Current Psychology*, vol. 33, no. 3, pp. 405–421, 2014.

[293] bHaptics Inc., “bhaptics tactsuit — the most advanced full-body haptic suit,” *Internet: https://www.bhaptics.com/tactsuit/ [March. 12, 2019]*, 2020.

[294] H. G. Hoffman, W. J. Meyer III, M. Ramirez, L. Roberts, E. J. Seibel, B. Atzori, S. R. Sharar, and D. R. Patterson, “Feasibility of articulated arm mounted oculus rift virtual reality goggles for adjunctive pain control during occupational therapy in pediatric burn patients,” *Cyberpsychology, Behavior, and Social Networking*, vol. 17, no. 6, pp. 397–401, 2014.

[295] G. N. Yannakakis and J. Togelius, “A panorama of artificial and computational intelligence in games,” *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 7, no. 4, pp. 317–335, 2014.

[296] J. Fürnkranz, “Machine learning in games: A survey,” *Machines that learn to play games*, pp. 11–59, 2001.

[297] T. Conde, W. Tambellini, and D. Thalmann, “Behavioral animation of autonomous virtual agents helped by reinforcement learning,” in *International Workshop on Intelligent Virtual Agents.* Springer, 2003, pp. 175–180.

[298] D.-W. Huang, G. Katz, J. Langsfeld, R. Gentili, and J. Reggia, “A virtual demonstrator environment for robot imitation learning,” in *2015 IEEE International Conference on Technologies for Practical Robot Applications (TePRA).* IEEE, 2015, pp. 1–6.

[299] S.-C. Yeh, M.-C. Huang, P.-C. Wang, T.-Y. Fang, M.-C. Su, P.-Y. Tsai, and A. Rizzo, “Machine learning-based assessment tool for imbalance and vestibular dysfunction with virtual reality rehabilitation system,” *Computer methods and programs in biomedicine*, vol. 116, no. 3, pp. 311–318, 2014.

[300] A. Borrego, J. Latorre, M. Alcañiz, and R. Llorens, “Comparison of oculus rift and

htc vive: feasibility for virtual reality-based exploration, navigation, exergaming, and rehabilitation," *Games for health journal*, vol. 7, no. 3, pp. 151–156, 2018.

[301] S. M. Palaniappan and B. S. Duerstock, "Developing rehabilitation practices using virtual reality exergaming," in *2018 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT).* IEEE, 2018, pp. 090–094.

[302] H. K. Kim, J. Park, Y. Choi, and M. Choe, "Virtual reality sickness questionnaire (vrsq): Motion sickness measurement index in a virtual reality environment," *Applied ergonomics*, vol. 69, pp. 66–73, 2018.

[303] T. Zhang, Z. McCarthy, O. Jow, D. Lee, X. Chen, K. Goldberg, and P. Abbeel, "Deep imitation learning for complex manipulation tasks from virtual reality teleoperation," in *2018 IEEE International Conference on Robotics and Automation (ICRA).* IEEE, 2018, pp. 1–8.

[304] I. Kastanis and M. Slater, "Reinforcement learning utilizes proxemics: An avatar learns to manipulate the position of people in immersive virtual reality," *ACM Transactions on Applied Perception (TAP)*, vol. 9, no. 1, pp. 1–15, 2012.

[305] A. Rovira and M. Slater, "Reinforcement learning as a tool to make people move to a specific location in immersive virtual reality," *International Journal of Human-Computer Studies*, vol. 98, pp. 89–94, 2017.

[306] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.

[307] M. Lanham, *Learn Unity ML-Agents–Fundamentals of Unity Machine Learning:*

Incorporate new powerful ML algorithms such as Deep Reinforcement Learning for games. Packt Publishing Ltd, 2018.

[308] A. Juliani, V.-P. Berges, E. Vckay, Y. Gao, H. Henry, M. Mattar, and D. Lange, "Unity: A general platform for intelligent agents," *arXiv preprint arXiv:1809.02627*, 2018.

[309] 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.

[310] 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.

[311] 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.

[312] 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.

[313] 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.

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

[315] 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.

[316] 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.

[317] 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.

[318] 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.

[319] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.

[320] J. Ho and S. Ermon, "Generative adversarial imitation learning," in *Advances in neural information processing systems*, 2016, pp. 4565–4573.

[321] 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.

[322] J. Wosik, M. Fudim, B. Cameron, Z. F. Gellad, A. Cho, D. Phinney, S. Curtis, M. Roman, E. G. Poon, J. Ferranti, J. N. Katz, and J. Tcheng, "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.

[323] 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.

[324] 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.

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

[326] 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.

[327] 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.

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

[329] 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.

[330] J. Chapleau, F. Canet, Y. Petit, G.-Y. Laflamme, and D. M. Rouleau, “Validity of goniometric elbow measurements: comparative study with a radiographic method,” *Clinical Orthopaedics and Related Research®*, vol. 469, no. 11, pp. 3134–3140, 2011.

[331] M. J. Kolber and W. J. Hanney, “The reliability and concurrent validity of shoulder mobility measurements using a digital inclinometer and goniometer: a technical report,” *International journal of sports physical therapy*, vol. 7, no. 3, p. 306, 2012.

[332] A. M. Bovens, M. A. van Baak, J. G. Vrencken, J. A. Wijnen, and F. T. Verstappen, “Variability and reliability of joint measurements,” *The American Journal of Sports Medicine*, vol. 18, no. 1, pp. 58–63, 1990.

[333] J. E. Giphart, J. P. Brunkhorst, N. H. Horn, K. B. Shelburne, M. R. Torry, and P. J. Millett, “Effect of plane of arm elevation on glenohumeral kinematics: a normative biplane fluoroscopy study,” *JBJS*, vol. 95, no. 3, pp. 238–245, 2013.

[334] K. Hayes, J. R. Walton, Z. L. Szomor, and G. A. Murrell, “Reliability of five methods for assessing shoulder range of motion,” *Australian Journal of Physiotherapy*, vol. 47, no. 4, pp. 289–294, 2001.

[335] D. L. Riddle, J. M. Rothstein, and R. L. Lamb, “Goniometric reliability in a clinical setting: shoulder measurements,” *Physical therapy*, vol. 67, no. 5, pp. 668–673, 1987.

[336] R. J. van de Pol, E. van Trijffel, and C. Lucas, “Inter-rater reliability for measurement of passive physiological range of motion of upper extremity joints is better

if instruments are used: a systematic review," *Journal of Physiotherapy*, vol. 56, no. 1, pp. 7–17, 2010.

[337] M. J. Tanaka, L. S. Oh, S. D. Martin, and E. M. Berkson, "Telemedicine in the era of covid-19: the virtual orthopaedic examination," *The Journal of bone and joint surgery. American volume*, 2020.

[338] R. R. Russo, M. B. Burn, S. K. Ismaily, B. J. Gerrie, S. Han, J. Alexander, C. Lenherr, P. C. Noble, J. D. Harris, and P. C. McCulloch, "Is digital photography an accurate and precise method for measuring range of motion of the shoulder and elbow?" *Journal of Orthopaedic Science*, vol. 23, no. 2, pp. 310–315, 2018.

[339] R. R. Russo, M. B. Burn, S. K. Ismaily, B. J. Gerrie, S. Han, J. Alexander, C. Lenherr, P. C. Noble, J. D. Harris, and P. McCulloch, "Is digital photography an accurate and precise method for measuring range of motion of the hip and knee?" *Journal of experimental orthopaedics*, vol. 4, no. 1, pp. 1–8, 2017.

[340] T. Russell, G. Jull, and R. Wootton, "Can the internet be used as a medium to evaluate knee angle?" *Manual therapy*, vol. 8, no. 4, pp. 242–246, 2003.

[341] P. A. Dent Jr, B. Wilke, S. Terkonda, I. Luther, and G. G. Shi, "Validation of teleconference-based goniometry for measuring elbow joint range of motion," *Cureus*, vol. 12, no. 2, 2020.

[342] M. Gholami, A. Ejupi, A. Rezaei, A. Ferrone, and C. Menon, "Estimation of knee joint angle using a fabric-based strain sensor and machine learning: A preliminary investigation," in *2018 7th IEEE International Conference on Biomedical Robotics and Biomechatronics (Biorob)*. IEEE, 2018, pp. 589–594.

[343] Y. Mengüç, Y.-L. Park, E. Martinez-Villalpando, P. Aubin, M. Zisook, L. Stirling, R. J. Wood, and C. J. Walsh, "Soft wearable motion sensing suit for lower limb biomechanics measurements," in *2013 IEEE International Conference on Robotics and Automation*. IEEE, 2013, pp. 5309–5316.

[344] J. H. Bergmann, S. Anastasova-Ivanova, I. Spulber, V. Gulati, P. Georgiou, and A. McGregor, "An attachable clothing sensor system for measuring knee joint angles," *IEEE Sensors Journal*, vol. 13, no. 10, pp. 4090–4097, 2013.

[345] H. Dejnabadi, B. M. Jolles, and K. Aminian, "A new approach to accurate measurement of uniaxial joint angles based on a combination of accelerometers and gyroscopes," *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 8, pp. 1478–1484, 2005.

[346] R. Garimella, T. Peeters, K. Beyers, S. Truijen, T. Huysmans, and S. Verwulgen, "Capturing joint angles of the off-site human body," in *2018 IEEE SENSORS*. IEEE, 2018, pp. 1–4.

[347] 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.

[348] N. Sarafianos, B. Boteanu, B. Ionescu, and I. A. Kakadiaris, "3d human pose estimation: A review of the literature and analysis of covariates," *Computer Vision and Image Understanding*, vol. 152, pp. 1–20, 2016.

[349] American Physical Therapy Association. (2019) Telehealth. [Online]. Avail-

able: https://www.apta.org/apta-and-you/leadership-and-governance/policies/telehealth

[350] 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.

[351] 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.

[352] 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/

[353] 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.

[354] 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.

[355] 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.

[356] 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.

[357] F. Doshi-Velez and B. Kim, "Towards a rigorous science of interpretable machine learning," *arXiv:1702.08608*, 2017. [Online]. Available: https://arxiv.org/pdf/1702.08608.pdf

[358] 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.

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

[360] Y. Yao, Z. Xiao, B. Wang, B. Viswanath, H. Zheng, and B. Y. Zhao, "Complexity vs. performance: empirical analysis of machine learning as a service," in *Proceedings of the 2017 Internet Measurement Conference*, 2017, pp. 384–397.

[361] 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.

[362] 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.

[363] 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.

[364] 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.

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

[366] M. R. Segal, “Machine learning benchmarks and random forest regression,” 2004.

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

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

[369] 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.

[370] 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.

[371] 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.

[372] 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.

[373] 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.

[374] D. Rand, R. Kizony, and P. T. L. Weiss, "The sony playstation ii eyetoy: low-cost virtual reality for use in rehabilitation," *Journal of neurologic physical therapy*, vol. 32, no. 4, pp. 155–163, 2008.

[375] B. Lange, C.-Y. Chang, E. Suma, B. Newman, A. S. Rizzo, and M. Bolas, "Development and evaluation of low cost game-based balance rehabilitation tool using the microsoft kinect sensor," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society.* IEEE, 2011, pp. 1831–1834.

[376] L. Rebenitsch and C. Owen, "Review on cybersickness in applications and visual displays," *Virtual Reality*, vol. 20, no. 2, pp. 101–125, 2016.

[377] S. N. Gieser, E. Becker, and F. Makedon, "Using cave in physical rehabilitation exercises for rheumatoid arthritis," in *Proceedings of the 6th International Conference on PErvasive Technologies Related to Assistive Environments.* ACM, 2013, p. 30.

[378] S. Finkelstein, A. Nickel, T. Barnes, and E. A. Suma, "Astrojumper: motivating

children with autism to exercise using a vr game," in *CHI'10 Extended Abstracts on Human Factors in Computing Systems.* ACM, 2010, pp. 4189–4194.

[379] H. Creagh, "Cave automatic virtual environment," in *Electrical Insulation Conference and Electrical Manufacturing & Coil Winding Technology Conference, 2003. Proceedings.* IEEE, 2003, pp. 499–504.

[380] T. Hatada, H. Sakata, and H. Kusaka, "Psychophysical analysis of the "sensation of reality" induced by a visual wide-field display," *Smpte Journal*, vol. 89, no. 8, pp. 560–569, 1980.

[381] M. C. Juan and D. Pérez, "Comparison of the levels of presence and anxiety in an acrophobic environment viewed via hmd or cave," *Presence: Teleoperators and virtual environments*, vol. 18, no. 3, pp. 232–248, 2009.

[382] 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.

[383] 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.

[384] 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.

[385] 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.

[386] 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.

[387] E. Cuervo, K. Chintalapudi, and M. Kotaru, "Creating the perfect illusion: What will it take to create life-like virtual reality headsets?" in *Proceedings of the 19th International Workshop on Mobile Computing Systems & Applications.* ACM, 2018, pp. 7–12.

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

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

[390] Neulog, "Gsr logger sensor nul-217," Neulog, Ed., 29.06.2019. [Online]. Available: https://neulog.com/gsr/

[391] 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.

[392] 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.

[393] 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.

[394] 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.

[395] I. NaturalPoint, "Naturalpoint, inc. dba optitrack motion capture system," 29.01.2019. [Online]. Available: https://optitrack.com/

[396] H. Corporation, "Htc corporation: Vive pro hmd," 29.01.2019. [Online]. Available: https://www.vive.com/us/product/vive-pro/

[397] 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.

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

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

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

[401] 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.

[402] 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.

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

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

[405] 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.

[406] 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.

[407] “COVID-19 pandemic,” Jun. 2021, page Version ID: 1027893345. [Online]. Available: https://en.wikipedia.org/w/index.php?title=COVID-19_pandemic&oldid=1027893345

[408] S. J. Yonter, K. Alter, M. N. Bartels, J. F. Bean, M. B. Brodsky, M. González-Fernández, D. K. Henderson, H. Hoenig, H. Russell, D. M.

Needham, S. Kumble, and L. Chan, “What Now for Rehabilitation Specialists? Coronavirus Disease 2019 Questions and Answers,” *Archives of Physical Medicine and Rehabilitation*, vol. 101, no. 12, pp. 2233–2242, Dec. 2020. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0003999320309369

[409] S. Ndumbe-Eyoh, P. Muzumdar, C. Betker, and D. Oickle, “‘Back to better’: amplifying health equity, and determinants of health perspectives during the COVID-19 pandemic,” *Global Health Promotion*, p. 175797592110009, Mar. 2021. [Online]. Available: http://journals.sagepub.com/doi/10.1177/17579759211000975

[410] “Position on Telehealth,” Sep. 2019. [Online]. Available: https://www.apta.org/apta-and-you/leadership-and-governance/policies/telehealth

[411] G. R. VandenBos and S. Williams, “The Internet versus the telephone: What is telehealth anyway?” *Professional Psychology: Research and Practice*, vol. 31, no. 5, pp. 490–492, 2000. [Online]. Available: http://doi.apa.org/getdoi.cfm?doi=10.1037/0735-7028.31.5.490

[412] E. Axelsson, E. Andersson, B. Ljótsson, D. Björkander, M. Hedman-Lagerlöf, and E. Hedman-Lagerlöf, “Effect of Internet vs Face-to-Face Cognitive Behavior Therapy for Health Anxiety: A Randomized Noninferiority Clinical Trial,” *JAMA Psychiatry*, vol. 77, no. 9, p. 915, Sep. 2020. [Online]. Available: https://jamanetwork.com/journals/jamapsychiatry/fullarticle/2765960

[413] M. E. Reed, J. Huang, I. Graetz, C. Lee, E. Muelly, C. Kennedy, and E. Kim, “Patient Characteristics Associated With Choosing a Telemedicine Visit vs Office Visit With the Same Primary Care Clinicians,” *JAMA*

Network Open, vol. 3, no. 6, p. e205873, Jun. 2020. [Online]. Available: https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2767244

[414] D. K. Shaw, "Overview of telehealth and its application to cardiopulmonary physical therapy," *Cardiopulmonary Physical Therapy Journal*, vol. 20, no. 2, pp. 13–18, Jun. 2009.

[415] A. C. Smith, E. Thomas, C. L. Snoswell, H. Haydon, A. Mehrotra, J. Clemensen, and L. J. Caffery, "Telehealth for global emergencies: Implications for coronavirus disease 2019 (covid-19)," *Journal of telemedicine and telecare*, vol. 26, no. 5, pp. 309–313, 2020.

[416] 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*, vol. 24, no. 5, p. 381, 2020.

[417] 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.

[418] 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.

[419] 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

[420] S. C. Gobron, N. Zannini, N. Wenk, C. Schmitt, Y. Charrotton, A. Fauquex,

M. Lauria, F. Degache, and R. Frischknecht, “Serious games for rehabilitation using head-mounted display and haptic devices,” in *International Conference on Augmented and Virtual Reality.* Springer, 2015, pp. 199–219.

[421] E. P. Cherniack, “Not just fun and games: applications of virtual reality in the identification and rehabilitation of cognitive disorders of the elderly,” *Disability and rehabilitation: Assistive technology*, vol. 6, no. 4, pp. 283–289, 2011.

[422] S. Finkelstein, A. Nickel, Z. Lipps, T. Barnes, Z. Wartell, and E. A. Suma, “Astrojumper: Motivating exercise with an immersive virtual reality exergame,” *Presence: Teleoperators and Virtual Environments*, vol. 20, no. 1, pp. 78–92, 2011.

[423] D. Deutscher, S. D. Horn, R. Dickstein, D. L. Hart, R. J. Smout, M. Gutvirtz, and I. Ariel, “Associations between treatment processes, patient characteristics, and outcomes in outpatient physical therapy practice,” *Archives of physical medicine and rehabilitation*, vol. 90, no. 8, pp. 1349–1363, 2009.

[424] R. S. Calabrò, A. Naro, M. Russo, A. Leo, R. De Luca, T. Balletta, A. Buda, G. La Rosa, A. Bramanti, and P. Bramanti, “The role of virtual reality in improving motor performance as revealed by eeg: a randomized clinical trial,” *Journal of neuroengineering and rehabilitation*, vol. 14, no. 1, p. 53, 2017.

[425] M. D. Wiederhold, M. Crisci, V. Patel, M. Nonaka, and B. K. Wiederhold, “Physiological monitoring during augmented reality exercise confirms advantages to health and well-being,” *Cyberpsychology, Behavior, and Social Networking*, vol. 22, no. 2, pp. 122–126, 2019.

[426] U. Qidwai, M. Ajimsha, and M. Shakir, “The role of eeg and emg combined

virtual reality gaming system in facial palsy rehabilitation-a case report," *Journal of bodywork and movement therapies*, vol. 23, no. 2, pp. 425–431, 2019.

[427] S. Ahn and S. Hwang, "Virtual rehabilitation of upper extremity function and independence for stoke: a meta-analysis," *Journal of exercise rehabilitation*, vol. 15, no. 3, p. 358, 2019.

[428] A. I. Corregidor-Sánchez, A. Segura-Fragoso, M. Rodríguez-Hernández, J. J. Criado-Alvarez, G.-G. Jaime, and B. Polonio-López, "Can exergames contribute to improving walking capacity in older adults? a systematic review and meta-analysis," *Maturitas*, 2019.

[429] M. C. Howard, "A meta-analysis and systematic literature review of virtual reality rehabilitation programs," *Computers in Human Behavior*, vol. 70, pp. 317–327, 2017.

[430] I. Brunner, J. S. Skouen, H. Hofstad, J. Aßmus, F. Becker, A.-M. Sanders, H. Pallesen, L. Q. Kristensen, M. Michielsen, L. Thijs *et al.*, "Virtual reality training for upper extremity in subacute stroke (virtues): a multicenter rct," *Neurology*, vol. 89, no. 24, pp. 2413–2421, 2017.

[431] S. R. S. BPT and N. Giladi, "Improved mobility and reduced fall risk in older adults after five weeks of virtual reality training," *Journal of Alternative Medicine Research*, vol. 9, no. 2, p. 171, 2017.

[432] P. Standen, K. Threapleton, A. Richardson, L. Connell, D. Brown, S. Battersby, F. Platts, and A. Burton, "A low cost virtual reality system for home based

rehabilitation of the arm following stroke: a randomised controlled feasibility trial," *Clinical rehabilitation*, vol. 31, no. 3, pp. 340–350, 2017.

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

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

[435] P. Rego, P. M. Moreira, and L. P. Reis, "Serious games for rehabilitation: A survey and a classification towards a taxonomy," in *5th Iberian conference on information systems and technologies.* IEEE, 2010, pp. 1–6.

[436] A. Elor and S. Kurniawan, "Deep Reinforcement Learning in Immersive Virtual Reality Exergame for Agent Movement Guidance," in *2020 IEEE 8th International Conference on Serious Games and Applications for Health (SeGAH).* Vancouver, BC, Canada: IEEE, Aug. 2020, pp. 1–7. [Online]. Available: https://ieeexplore.ieee.org/document/9201901/

[437] E. L. Deci and R. M. Ryan, "Self-determination theory." 2012.

[438] L. A. Shaw, B. C. Wünsche, C. Lutteroth, S. Marks, J. Buckley, and P. Corballis, "Development and evaluation of an exercycle game using immersive technologies," 2015.

[439] L. A. Shaw, J. Buckley, P. M. Corballis, C. Lutteroth, and B. C. Wünsche, "Competition and cooperation with virtual players in an exergame," *PeerJ Computer Science*, vol. 2, p. e92, 2016.

[440] A. Keesing, M. Ooi, O. Wu, X. Ye, L. Shaw, and B. C. Wünsche, "Hiit with hits: Using music and gameplay to induce hiit in exergames," in *Proceedings of the Australasian Computer Science Week Multiconference*, 2019, pp. 1–10.

[441] L. A. Shaw, B. C. Wuensche, C. Lutteroth, J. Buckley, and P. Corballis, "Evaluating sensory feedback for immersion in exergames," in *Proceedings of the Australasian Computer Science Week Multiconference*, 2017, pp. 1–6.

[442] J. C. Haller, Y. H. Jang, J. Haller, L. Shaw, and B. C. Wünsche, "Hiit the road: Using virtual spectator feedback in hiit-based exergaming," in *Proceedings of the Australasian Computer Science Week Multiconference*, 2019, pp. 1–9.

[443] C. D. Batson, P. A. Van Lange, N. Ahmad, and D. A. Lishner, "Altruism and helping behavior," *The Sage handbook of social psychology*, vol. 1, pp. 279–295, 2003.

[444] M. O. Wilhelm and R. Bekkers, "Helping behavior, dispositional empathic concern, and the principle of care," *Social Psychology Quarterly*, vol. 73, no. 1, pp. 11–32, 2010.

[445] K. Doherty and G. Doherty, "Engagement in hci: conception, theory and measurement," *ACM Computing Surveys (CSUR)*, vol. 51, no. 5, pp. 1–39, 2018.

[446] J. A. Russell, "A circumplex model of affect." *Journal of personality and social psychology*, vol. 39, no. 6, p. 1161, 1980.

[447] P. Ekman, "Basic emotions," *Handbook of cognition and emotion*, vol. 98, no. 45-60, p. 16, 1999.

[448] 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.

[449] 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.