UNIVERSITY OF CALIFORNIA,
IRVINE

Engineering for Arm Use After Stroke: A Precision Rehabilitation Model, Minimalistic Robot Design Pattern, and Proprioception-Targeting Gaming Paradigm

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Mechanical and Aerospace Engineering

by

Dylan Scott Reinsdorf

Dissertation Committee:
Professor David J. Reinkensmeyer, Chair
Assistant Professor Alexandra Voloshina
Professor Xiaohui Xie

2022

DEDICATION

To my parents for your love, support, and enthusiasm for my graduate studies.

To Liz, the love of my life, my dearest friend, and most trusted companion, thank you for joining me on this journey. You brought care, support, and love during times when I needed it most. From the bottom of my heart, thank you.

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

Page

ACKNOWLEDGEMENTS

To my advisor Dr. David Reinkensmeyer, thank you for the opportunity to work on such challenging problems with you. This Ph.D. experience has truly become all that I had hoped for, and I credit a lot of that to you. You gave me opportunities in so many ways: open-ended problems, minimal oversight, and the ability to steer my own research - all opportunities that I sought, but never experienced at such magnitude. You fostered creativity and exploration always, and just made research a joy. I will miss your constant interest and curiosity. And somehow, you maintained a steady, positive environment through a shifting background of pandemic lock downs and tragic current events. Thank you, Dave. To Dr. Sasha Voloshina and Dr. Xiaohui Xie, thank you for your support and feedback throughout this process.

Thank you to my current and past Biorobotics lab mates, as much of this research was built on and with your contributions. Erin, you were truly a joy to work with. You taught me how to be a better mentor, and I so look forward to seeing how your own Ph.D. and career progress. Luis, Daniel, Andria, and Caitlin, thank you for continuing with and carrying this research forward. Luis and Daniel thank you for your immense data collection efforts. Andria thank you for jumping in the deep end, taking the reins from me, and taking the time to understand every detail of this work so we can study it further in stroke survivors. Veronica, thank you for being such a thoughtful, interested, and truly helpful lab mate – studying and practicing both engineering and pedagogy with you was a joy. Chris, thank you for the many meandering conversations about robotics, biomechanics, proprioception, study designs, and just learning how to do research. Continuing to connect with you and share research ideas through the pandemic enriched my Ph.D. experience, thank you. To Martí, thank you for your expertise, you are truly an amazing designer. To Quentin, Merav, Shusuke, Juan, Dongwon, and Han Gil thank you for sharing your own experiences and research with me, always listening to and providing insightful feedback. To Vicky, thank you for your clinical expertise and for passing on a small sliver of how to work with stroke survivors, I would have been lost without you.

I would be remiss if I did not thank Anthony Anderson, my lab mate during my MS research and friend. Anthony, thank you for all your support in helping me transition into a novel space and into academic research itself. As with all the wonderful people who I had the privilege of sharing my graduate research experience with, our paths did not separate, they are growing in interesting new directions that we can share with one another.

CURRICULUM VITAE

EDUCATION

University of California, Irvine

Ph.D. Mechanical Engineering — Summer 2022

Thesis: Sharpening Our Sixth Sense: A Simple Robot Design Pattern and Gamified Training Paradigm for Improving Proprioception, Toward Increasing Daily Arm Use After Stroke

Courses: Human Sensory Motor Systems, Biorobotics, Artificial Intelligence, Advanced Pedagogy

University of Washington, Seattle

M.S. Mechanical Engineering — June 2019

Thesis: Development of a Robotic Unloader Brace for Investigation of Conservative Treatment of Medial Knee Osteoarthritis

Courses: Probabilistic Robotics, Linear Systems Theory, Linear Multivariable Control, Digital Control Systems Design, Fundamentals of Optimization, Numerical Optimization, Parallel Computing, Applied Mathematics

California Polytechnic University, San Luis Obispo — Overall GPA 3.3

B.S. Mechanical Engineering — June 2012

Capstone project: Quantifying Proximal Conformability of a Thoracic Aortic Stent Graft

Technical elective courses: Finite Element Analysis, System Dynamics, Advanced Mechanics of Materials, Electromagnetics

RESEARCH EXPERIENCE

UC Irvine Biorobotics Lab

Graduate Student Researcher — Fall 2019 – Summer 2022

Leading projects to investigate the effects of gamified finger proprioception training on motor learning after stroke, and to implement improved clinical assessments for finger extension and finger proprioception. Developed novel gamified training paradigm and compact robotic device to enhance proprioceptive learning and investigated efficacy in human subjects testing.

Center for Limb Loss and Mobility

VA Puget Sound, Graduate Research Assistant — Fall 2017 - Summer 2019

Led project with VA Hospital clinicians to discover an improved treatment method for knee osteoarthritis using a robotic unloader brace. Ideated brace concept based on deficiencies in braces and current treatment options. Developed working prototype of robotic brace with custom knee unloading modulation mechanism and off-board actuation system for real-time brace control. Evaluated performance in human subjects testing.

INDUSTRY EXPERIENCE

Elixir Medical

Process Development Engineer — October 2016 – August 2017

Independently designed, built, and tested plasma treatment robot to control catheter manufacturing process. Optimized manufacturing processes for sophisticated FDA and CE regulated medical device (drug eluting bioresorbable coronary stent) to improve clinical device performance and quality.

Manufacturing Engineer — October 2014 – October 2016

Designed and executed experiments to develop design and process improvements for cardiovascular stent. Led R&D, Operations, Regulatory Affairs, and Quality Assurance teams to obtain regulatory approval for and implement product development projects. Independently diagnosed, contained, and implemented solutions for line down issues at overseas contract manufacturing facility.

Covidien, GI Solutions

Associate Manufacturing Engineer — August 2012 – October 2014

Led process development activity for novel esophageal cell collection device. Assisted cross functional team in device design, risk analysis, and clinical trial activities. Patents listed in §Patents below. Led eight-person R&D, Operations, Regulatory Affairs, and Quality Assurance team to implement Covidien product brand across all current products. Received international regulatory (FDA, CE, etc.) approval.

Proof of Concept, LLC

Mechatronics Engineer — June 2011 – June 2012

Interfaced directly with clients to develop medical device and consumer product solutions: motorized bike rack for local entrepreneur and semi-automated manufacturing equipment for EV3 Neurovascular, as an independent contributor in small, high energy, fast paced engineering consulting firm.

C.R. Bard, Davol Inc.

Mechanical Engineering Intern — Summer 2010

Designed tooling for hernia fixation device assembly and led risk analysis for new automated equipment.

PUBLICATIONS

1. D. Reinsdorf, E. Mahan, D. Reinkensmeyer. Proprioceptive Gaming: Making Finger Sensation Training Intense and Engaging with the P-Pong Game and PINKIE Robot, 2021 IEEE 43rd Engineering in Medicine and Biology Conference.
2. L. Wu, E. Zhu, C. Callaghan, D. Irwin, D. Reinsdorf, V. Swanson, A. Zwirn, D. Reinkensmeyer. Rapidly Converting a Project-Based Engineering Experience for Remote Learning: Successes and Limitations of Using Experimental Kits and a Multiplayer Online Game. Advances in Engineering Education, 2020.
3. D. Reinsdorf, C. Richburg, J. Czerniecki, P. Aubin, Development of a Robotic Unloader Brace for Investigation of Conservative Treatment of Medial Knee Osteoarthritis, 2019 IEEE 16th International Conference on Rehabilitation Robotics.

PATENTS

1. Lubinski, A., Maguire, M. and Reinsdorf, D. (2016). Regions of varying physical properties in a compressible cell collection device. US 20160081677 A1. Issued March 24, 2016.
2. Lubinski, A., Maguire, M. and Reinsdorf, D. (2016). Scored Retaining Features in a Compressible Cell Collection Device. US 20160081670 A1. Issued March 24, 2016.
3. Lubinski, A., Maguire, M. and Reinsdorf, D. (2016). Use of Expansion-Force Elements in a Compressible Cell Collection Device. US 20160081671 A1. Issued March 24, 2016.

TEACHING EXPERIENCE

Mechanical Systems Laboratory, Teaching Assistant — Spring 2020
UC Irvine Department of Mechanical and Aerospace Engineering

Time and Frequency Response of Second Order Systems, Guest Lecturer — May 13, 15 20

AWARDS

Pedagogical Fellow, UCI Division of Teaching Excellence and Innovation — December 2020

Minority Serving Institution Fellowship, UCI Graduate Division — September 2019

Predoctoral Training Grant, VA Puget Sound Healthcare System — November 2018

Conference Travel Grant, University of Washington, Seattle Graduate School — July 2018

National Travel Grant, Center for Neurotechnology, University of Washington — June 2018

2018 Fellow, Amplifying Human Movement and Performance Lab, UW — March 2018

TALKS

1. "**Robots and Sensors – Technologies for Stroke Rehabilitation,"** at the 2022 UCI Chief Executive Winter Roundtable on March 22, 2022.
2. "**Proprioceptive Gaming: Making Finger Sensation Training Intense and Engaging with the P-Pong Game and PINKIE Robot,"** at the 2021 IEEE 43rd Engineering in Medicine and Biology Conference.
3. **"Development of a Robotic Unloader Brace for Investigation of Conservative Treatment of Medial Knee Osteoarthritis,"** at the International Conference on Rehabilitation Robotics in Toronto, ON, Canada on June 27, 2019.
4. "**Performance Evaluation of a Robotic Unloader Brace for Medial Knee Osteoarthritis Treatment Investigation,"** at the Northwest Biomechanics Symposium in Bozeman, MT on May 18, 2019.
5. **"Evaluation of an Active Unloader Brace for Medial Knee Osteoarthritis,"** at the Center for Neurotechnology Rextravaganza in Seattle, WA on November 26, 2018.
6. **"System Identification of an Active Unloader for Medial Knee Osteoarthritis,"** at the VA Puget Sound Healthcare System Young Investigator Symposium in Seattle, WA on September 20, 2018.
7. **"Evaluation of an Active Unloader Brace for Medial Knee Osteoarthritis,"** at the American Society of Biomechanics Annual Meeting in Rochester, RN on August 11, 2018.

ABSTRACT OF THE DISSERTATION

Engineering for Arm Use After Stroke: A Precision Rehabilitation Model, Minimalistic Robot Design Pattern, and Proprioception-Targeting Gaming Paradigm

by

Dylan Scott Reinsdorf

Doctor of Philosophy in Mechanical and Aerospace Engineering

University of California, Irvine, 2022

Professor David J. Reinkensmeyer, Chair

An estimated one in four people will experience a stroke in their lifetime. One of the most debilitating and common consequences of a stroke is loss of sensorimotor function on one side of the body. In this dissertation we pose the question: what should we target as we develop robotic and sensor-based tools to increase paretic upper extremity use after stroke? We approach this question by identifying three gaps. First, we lack understanding of how impairment reduction can lead to use increase. Second, despite the prevalence of proprioceptive deficits after stroke and the potential role of proprioception in motor learning, there are no methods for intensely and engagingly training hand propriomotor capacity. Third, there is an unmet need for compact rehabilitation robotic devices suitable for home use. To address these gaps, this dissertation presents advances in precision medicine, rehabilitation gaming paradigms, and rehabilitation robot design.

We identify responders to technology-based training by developing a model to explain changes in daily arm use after therapy, analyzing data from seven robotic clinical trials

conducted by our laboratory. The identified model demonstrated that individuals with low baseline use relative to their baseline score on a common clinical measure of hand dexterity (a mismatch that we call "untapped use potential") had a high probability of increasing use, independent of the type of study intervention.

But what of the non-responders? The model predicts that an increase in dexterity would help. We considered this finding in light of a previous study that found that finger proprioception acuity predicted participants' ability to change their dexterity after intense robotic movement training. The problem is that few paradigms exist for retraining finger proprioception acuity. Thus, we developed a novel proprioception-targeting gaming paradigm, Propriopixels, for simultaneously training finger motor function and proprioception. Instead of displaying all game elements on screen as in a traditional video game, in the Propriopixels paradigm one of the game dimensions is conveyed to the finger with a robotic device. That is, we create a "Propriopixel" by moving the finger instead of a light pixel on screen.

We then asked, "What is the minimal robot needed to implement the Propriopixel paradigm?" Compared to more common robotic therapy designs that utilize high-cost actuators and sensors to render a continuum of impedances, we propose the design concept of a binary impedance robot that only renders the two limits – high and low impedance. It is either stiff to passively drive the Propriopixel finger, or transparent to sense active finger movements for a game input. Given the savings that solely stiff, actuated and transparent, unactuated mechanisms afford, a Propriopixels game can be cost-effectively realized with a relatively simple, binary impedance robotic device, that we demonstrate with a device called PINKIE.

We implemented Propriopixels with Proprioceptive-Pong, a game based on the classic Atari arcade game. We used the PINKIE device to study a purely passive finger movement version of Proprioceptive-Pong and the FINGER robotic exoskeleton to train an active movement version, both with unimpaired participants. We found that training with the passive version of the game yielded gains in passive proprioception acuity, while training with the active version of the game yielded gains in active but not passive proprioception acuity, suggesting a specificity of proprioceptive training principle and/or important differences between the passive proprioceptive acuity assessments deployed on each robotic device.

Following this, we studied the extent to which stroke survivors could understand and play Proprioceptive-Pong. We evaluated two methods of controlling success rates during Proprioceptive-Pong play, by either adjusting virtual game parameters only or by physically assisting players to complete successful movements. We found that a progression of game modes that gradually grew in complexity was effective for teaching Proprioceptive-Pong, and that the success control algorithm was capable of regulating success with both methods of assisting participants - virtually and physically. These results indicate that stroke survivors can understand and play Proprioceptive-Pong.

CHAPTER 1: INTRODUCTION

An estimated one in four people worldwide will experience a stroke in their lifetime [1]. While approximately one third of strokes result in death [2], an estimated 75% of survivors experience difficulties completing activities of daily living [3]. One of the most common deficits after stroke is hemiparesis, the loss of volitional movement and weakness on one side of the body [4]. Hemiparesis has debilitating consequences: upper extremity impairments limit functional independence for an estimated 50% of stroke survivors [5], stemming from deficits in motor function and sensation which play integral roles in how we participate in activities and interact with our environment.

A. Upper extremity rehabilitation, from impairment to use

Rehabilitation is a major aspect of recovery for stroke survivors that occurs over several months after stroke. Rehabilitation can focus on different aspects of disability. Early rehabilitation during the acute phase of stroke tends to focus on body functions and structures, which describes actual anatomy and physiology of the body, and later toward the chronic stroke phase rehabilitation tends to focus on activity limitations, participation restrictions, and quality of life [6]. Each are separate dimensions of World Health Organization's International Classification of Functioning, Disability and Health [7], and little correlation has been found between body structure level and activity and participation level dimensions [6], [8]. This means that, although assessments exist to quantify each, we lack an understanding of how improvements at the physiological level map to improvements at the activity and participation level, likely due to the many factors upon which activity and participation depend [9], and that such factors may vary based on the unique health state, lifestyle, and goals of each stroke survivor. For example, in upper extremity rehabilitation, it has been found that reduction in upper extremity impairment does not necessarily translate to increased upper extremity use, and that in fact upper extremity use can decrease following therapy [10]. We identify this as the first of three gaps that this dissertation addresses. That is, presently in upper extremity stroke rehabilitation we lack mechanistic links between improvements in impairment or capacity to improvements in activity. We interpret this gap as a need to identify predictors, specifically what therapy target, to generate lasting gains in upper extremity use.

B. Upper extremity stroke rehabilitation technology

Given the large variation in the many factors that contribute to recovery and upper extremity use, the development of precision rehabilitation models that guide clinicians in designing a therapy program based on a patient's unique assessment will require a substantial amount of data, a problem well suited for robotic and sensor-based technology. Furthermore, considering the large number of practice repetitions required for sensorimotor learning [11]–[13] and the relatively few number practiced during therapy sessions [14], home-based therapy is a major aspect of stroke rehabilitation. Home-based therapy adherence has been shown to be low [15], highlighting another need technology is well poised to address. Over the past two decades there has been an increase in the use of mechatronic devices in upper extremity rehabilitation, and a growing interest in engineering devices for home use.

To our knowledge, only two commercial devices exist that provide real-time feedback of the upper extremity activity, the MiGO by Flint Rehab and the ARYS by Tyromotion. Although few commercial devices exist specific for arm activity tracking, in research the use of wrist and body worn accelerometers to measure and give feedback on arm use amount and quality is increasing [16], [17]. Not only are wearables well suited to objectively quantify real-world upper extremity use, their use as feedback devices to motivate exercise is also being investigated by multiple groups. In their 2017 systematic review of interactive wearable systems, Wang et al. identified three studies that presented sensor-based devices for motivating exercise with automated feedback [18].Two studies presented early-stage device design and engineering performance verification: Myllymaa et al. developed an automated haptic feedback of arm movement [19] and Jeong et al. developed an automated alarm system to monitor exercise levels with a stationary arm bike [20]. One study investigated the effects of automated feedback on the daily arm activity of stroke survivors [21]. Holden et al. used their novel wrist worn accelerometer automated feedback system "CueS" to cue participants to complete a rehabilitative arm movement every hour over a one week period and found that activity levels increased immediately following cues, and that participants reported increased overall upper extremity activity levels. This means that sensor-based devices can and are being used for much more than measurement alone. Commercial examples among them are the MusicGlove by Flint Rehab which is used to practice finger individuation with a Guitar-Hero like game, and FitMi also by Flint Rehab for practicing hand, arm, core, and leg exercises. Sensor-based devices offer attractive properties over robotic device in that they are typically less expensive, lighter weight, and safer, with the disadvantage of requiring the user to perform all movements actively, without assistance. In fact, in the study of the MusicGlove wearable grip sensor, Sanders et al. found that only 13% of stroke survivors admitted to a hospital without limiting complications had adequate hand function to use the device. So, while sensor-based devices have many advantages that make them feasible for implementation into home-based training, they require moderate levels of hand function to operate.

Rehabilitation robots offer distinct advantages of accommodating lower levels of hand function. Several commercially available upper extremity exoskeleton robots exist. In a recent systematic review, Gull et al. identified 16 devices, ten of which are shown in Figure 1 [22]. Among these devices six are designed for the hand, the Amadeo by Tyromotion, the MyoPro by Myomo which also actuates the elbow and wrist joints, the Hand of Hope by Rehab-Robotics, the HandTutor by MediTouch, SEM glove by Bioservo, and Exo Glove Poly (although we were not able to confirm that it is commercially available). Among these products, only the Hand of Hope, ArmTutor, and Amadeo are designed for rehabilitation and must be used under clinical supervision. And while some of these products are relatively compact and lightweight through the application of soft robotic mechanisms, they are not capable of actuating intricate finger individuation movements like their non-soft counterparts.

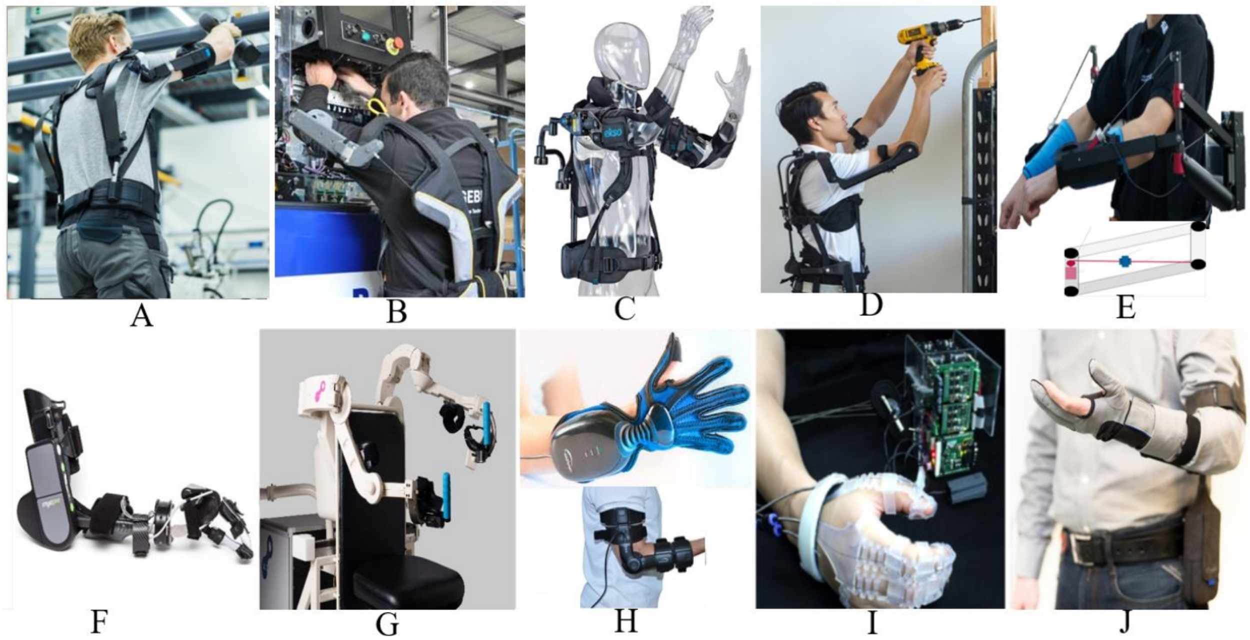

Figure 1. Commercially available upper extremity robots.

From left to right: Skelex (A), Egrosquelettes by GOBIO-robot (B), EksoVest by Ekso Bionics (C), Modular Agile eXoskeleton by SuitX (D), Robo-Mate (E), MyoPro Orthosis by Myomo, Inc. (F), Alex exoskeleton by Kinetek Wearable Robotics (G), Hand and Arm tutor by MediTouch (H) and SEM glove by BioServo (J). We were not able to confirm that The Exo Glove Poly (I) is commercially available.

In reviewing these devices we identify a second gap that this dissertation addresses: considering that most home-based therapy is a major aspect of stroke-rehabilitation and the range of hand motor deficits in stroke survivors, there is an unmet need for compact rehabilitation robotic devices capable of delivering both assessments and therapy that can be used in the home, without the supervision of a clinician. And, ideally, such home-based robots would be as simple-as-possible, meeting a “minimal design criteria” defined by a mechanistic understanding of what is necessary in home training to ultimately promote increased hand use.

C. Gamified sensorimotor rehabilitation

Gamified rehabilitation has been shown to be effective for facilitating high repetition hand movement practice in the home [23]. A common thread among these gamified solutions is that, borrowing from substantial advances in the commercial video game industry, they predominantly convey information visually. The specificity of learning hypothesis states that people learn by using the optimal source of information for achieving the goal of the activity [24]. In fact, vision has been found to dominate other afferents like proprioception [25]–[27] which may a be key input to motor learning [28], [29].

Somatosensory deficits in touch, pain, temperature, and proprioception have been reported to range from 11% to 100% after stroke [30]–[33]. This wide variability is likely due to the range of measures used to quantify somatosensory deficits, the challenge involved in developing reliable measures, and the wide umbrella of the multiple somatic sensations it encompasses. Proprioception and two-point touch discrimination have been identified as potentially the most important somatic sensations for hand function [34], [35].

In this dissertation, we will adopt a broad definition of proprioception as the sense of movement without vision. Narrower definitions have been proposed using proprioception and kinesthesia, with proprioception describing static position sense and kinesthesia describing motion sense, or with proprioception describing the non-conscious processing of proprioceptors and kinesthesia as the conscious processing of proprioceptors [36]. Here we will not make such a distinction. Of note, it is rare in studies involving passive movement to confirm the absence of muscle activations using electromyographic recordings [37], to our knowledge there is no assessment for non-conscious proprioception, and in studies that train proprioception it is often unclear to what extent conscious and non-conscious proprioception are being trained. The involvement of conscious proprioception is clear when training involves querying the participant about where their extremity is located or having participants verbally confirm when they have reached a target. However, in commonly studied active movement paradigms such as balancing, tai chi, and yoga, it is unclear as to the level of conscious versus non-conscious processing that is being trained. And as we do not have a way to isolate and measure the acuity of non-conscious proprioception, the specificity of training to each type of processing and the transfer of training conscious proprioception to non-conscious has to our knowledge, yet to be investigated. In summary, although there are neurophysiological motivations for adopting more precise definitions of proprioception, we will employ a broad definition in summarizing the state of proprioception assessment and training research.

Returning to gamified rehabilitation, one attractive property of games is the freedom to control the flow and quality of information, which is theoretically attractive for optimizing learning. This raises the question, do visually driven games allow stroke survivors to compensate for somatosensory deficits with vision during training, and thus reduce the potential benefits of gamified training? Should rehabilitation games limit the use of vision and balance the use of other afferents? Or perhaps more simply, how can rehabilitation games be designed to target proprioception training?

In reviewing the literature, we found only one example of a proprioception-targeting training game. We define a proprioception-targeting game as one that requires players to use proprioception to make in-game decisions, or put another way, a game that cannot be played by compensating with other senses like vision. In the gaming paradigm developed by

Elangovan et al. participants tilted a virtual table to move a ball into a hole using a wrist robot [38]. The ball position, not shown on screen, was encoded through vibration frequencies of vibrotactile motors placed on the participant's forearm. Participants closed their eyes during the task and the task was performed with two aspects of somatosensation: touch (the ball position) and active proprioception (table tilt position). They studied this paradigm with twelve stroke survivors and found that passive wrist proprioceptive acuity significantly improved from baseline to post and was retained at a two day follow up. Other proprioception training games have been developed for balance with instrumented balance boards, and certainly these involve proprioception during play as they require balance, however they do not explicitly isolate it as a sense [39], [40]. If we were to widen our definition to include games such as these, then any game that involves active movement could be considered proprioception-targeted training. In reviewing these games, we identify a third gap that this dissertation addresses: despite the prevalence of proprioceptive deficits after stroke and the potential role of proprioception in motor learning, there are no proprioception-targeting games to intensely and engagingly train hand propriomotor capacity. Further, building on our previously identified gap in hand rehabilitation robots, proprioception-targeted training is a well suited application for robotics: with robotics, proprioception can be truly isolated from motor function by passively moving appendages in the absence of user movement, which is attractive from a scientific perspective.

D. Summary of this dissertation

Thus, in our review of the field of robotic and sensor-based therapy design, we identified three key gaps. First, there is a need to identify predictors, specifically of who can benefit from technology-based training, and what such therapy should target, in terms of generating lasting gains in upper extremity use. Second, there is an unmet need for compact rehabilitation robotic devices capable of delivering both assessments and therapy that can be used in the home. Third, there are no proprioception-targeting games to intensely and engagingly train hand propriomotor capacity. Therefore, this dissertation focuses on making advances in three primary areas: precision rehabilitation (Chapter 2), the design of a novel class of rehabilitation games (Chapter 3-5), and the design of minimalistic rehabilitation robotics (Chapter 3). Next, we summarize the methods and findings of each chapter.

In Chapter 2 we identify proprioception as a therapy target by developing a model to explain changes in daily arm use after therapy. By analyzing data from seven stroke rehabilitation clinical trials conducted by our laboratory, we found that baseline measures of hand dexterity and the amount of daily hand use best explained whether patients increased daily arm use after therapy. Patients who increased arm use were mismatched in capacity and performance in a particular way: their dexterity was high relative to their low amount of use; patients with low dexterity, in contrast, seldom increased daily arm use. We considered this finding in light of a previous study that found that finger proprioception acuity predicted participants' ability to change their dexterity after intense robotic movement training [41]. Proprioception has also been found to predict increase in daily arm use one-year post stroke [42]. Mechanistically, it is reasonable to expect that finger proprioception is needed for dexterity recovery, since finger proprioception likely provides the teaching signal used by the brain to identify viable residual descending pathways during movement practice [29]. Taken together, this analysis suggests that training finger proprioception may hold the key for restoring daily arm use.

In Chapter 3 we present the development of a novel proprioception-targeting gaming paradigm, Propriopixels, for simultaneously training finger motor function and proprioception. Home-based therapy outcomes may be ameliorated by game-based therapy, which has been shown to increase motivation and the amount of practice achieved [43], [44]. While games have been developed that incorporate simultaneous sensory feedback (vision, touch, proprioception) of game elements [39], [40], we have found only one example of a game that requires proprioception to make gameplay decisions [38]. Instead of showing all game elements on screen as in a traditional video game, in the Propriopixels paradigm one of the game dimensions is conveyed to the finger with a robotic device. That is, we create a "Propriopixel" by moving the finger instead of a light pixel on screen.

Propriopixels gives rise to a simple robot design pattern, which we call a "binary impedance robot". Compared to more common robotic therapy designs that utilize high-cost actuators and sensors to render a continuum of impedances, a binary impedance robot only renders the two limits – high and low impedance. It is either stiff to passively drive the Propriopixel finger, or transparent to sense active finger movements for a game input. Given the savings that solely stiff, actuated and transparent, unactuated mechanisms afford, a Propriopixels game could be realized with a relatively simple, binary impedance robotic device.

We implemented Propriopixels with Proprioceptive-Pong, a game based on the classic Atari arcade game, and the binary impedance robot PINKIE, a practical, low-cost device that can be built with rapid manufacturing techniques and is capable of either driving or sensing movements of the index and middle fingers. In a pilot study of 15 unimpaired participants, we showed that playing Proprioceptive-Pong for 15 minutes significantly improved passive finger proprioceptive acuity, while playing a traditional video-only version of the game did not. Although the Proprioceptive-Pong group did not improve significantly more than the video-only group, the video-only group had significantly more repetitions and higher success than the Proprioceptive-Pong group, both of which have been shown to be beneficial for learning [45]–[47].

In Chapter 4 of this dissertation, we further investigate the potential benefits of Propriopixels training by quantifying training effects on passive proprioception acuity, activity proprioception acuity, and dexterity. We first present improvements to the Proprioceptive-Pong game derived from lessons learned in Chapter 3, namely novel stimuli in the form of targets, a new non-robotic proprioception-targeting training mode that we call "Visioception", leveling, and an automatic difficulty control algorithm. In keeping with our goal of developing simplified robotic technology, the algorithm regulated difficulty by adjusting virtual parameters only, without physically forcing the player. We studied the effects of three types of active movement training, the two proprioception-targeting paradigms Propriopixels and Visioception and a Video-only training group for which the game elements were always displayed on screen like a typical video game. All groups were matched in training success rates, repetitions, and time. Evaluating our Proprioceptive-Pong developments, we demonstrated that we were able to both regulate success completely virtually without the physical assistance capabilities of a complex robotic device and that targets drove learning: Propriopixels players began with initially high error on the targets and reduced that error over the course of play. We also found that all groups significantly decreased position errors during Proprioceptive-Pong play and that all groups significantly improved arm dexterity. However, there were no changes in passive proprioception acuity, which we found surprising for the Propriopixels group considering that their training repetitions and success rates increased, a finding which we further investigated.

Considering the complexity of the Proprioceptive-Pong game and our target stroke user population, in Chapter 5 we evaluated the extent to which stroke survivors could understand and play Proprioceptive-Pong that we developed in Chapters 3 and 4. In a pilot study of three stroke survivors, we found that a progression of game modes that gradually grew in complexity was effective for teaching Proprioceptive-Pong. By the end of the progression, participants were able to play a game that required sensing through multiple afferents, cognitively comparing the multimodal sensory information, and actively moving the index finger. Together these results indicate that stroke survivors can understand and play Proprioceptive-Pong, and that short bouts of play can cause gains in passive and active proprioceptive acuity. Further, Proprioceptive-Pong with virtual success control can be implemented with a simple binary impedance robot like PINKIE. Finally in Chapter 6, we review the major contributions of this work and discuss future research directions.

CHAPTER 2: WHAT SHOULD WE TARGET TO IMPROVE RECOVERY OF DAILY LIMB USE AFTER STROKE: A MODELING ANALYSIS OF ROBOT- AND SENSOR-BASED CLINICAL TRIALS

SUMMARY OF THE CHAPTER

Even with a substantial amount of functional recovery of the hand, most stroke patients report difficulty using their limb in daily life, a finding supported by recent wearable sensing studies. Improving rehabilitation outcomes is current limited by a lack of understanding of the key ingredients to restoring daily arm use. In this retrospective study, we analyzed data from seven upper extremity stroke rehabilitation clinical trials that administered novel, technology-based therapeutic interventions to identify independent variables that explained responsiveness to intervention, in terms of increasing self-reported daily arm use. Specifically, we formulated the dependent variable as a binary indicator of increase in daily arm as reported on the Motor Activity Log Amount of Use subscale score comparing baseline (before therapy) and follow up (1-3 months after therapy). We performed variable selection on data for individuals in the chronic stroke stage of stroke to build a multiple logistic regression model. We then evaluated model sensitivity and stability using bootstrap resampling-based methodology and model generalizability by testing the same model structure with data from individuals in the subacute stroke phase, using the independent variables time after stroke and those selected using the chronic stroke phase analysis. The final model contained two independent variables measured at baseline, the baseline Box and Blocks Test z-score and the Motor Activity Log Amount of Use rating. A one unit increase in Box and Blocks Test z-score raised the odds of increasing daily arm use after therapeutic intervention by 161%, while a one unit increase in baseline Motor Activity Log Amount of Use reduced the odds by 50%. This model was stable across 10,000 bootstrap resamples –

both independent variables were included in selected models most often (99.9% and 88.5% of models for the Box and Blocks Test and Motor Activity Log Amount of Use subscale, respectively), and the model itself was selected most often (8.6% of bootstrap resamples). For people in the subacute stroke phase, there was a non-significant relationship between both independent variables and the dependent variable ($p = 0.21$, $p = 0.76$ for the Box and Blocks Test and Motor Activity Log Amount of Use subscale, respectively).

The identified model highlights a subject-specific predictor of the practical efficacy of robot-assisted rehabilitation: participants with a mismatch in baseline amount of arm use and baseline dexterity (i.e. participants with low Motor Activity Log Amount of Use relative to Box and Blocks Test z-score) tended to increase use. We call this situation "untapped use potential", and it predicted an uptake of use independent of the type of study intervention, technology-based or conventional. This relationship, however, did not hold in a small sample of subacute participants. Instead, an indicator of spontaneous recovery, time after stroke, explained their increase in daily arm use.

INTRODUCTION

Stroke is a leading cause of long-term disability [48]. Globally, one in four adults will have a stroke in their lifetime [49], and stroke causes a death every four minutes [48]. Not only is stroke mortality high, an estimated 50% of stroke survivors are chronically disabled [49], [50].

The aim of stroke rehabilitation is to improve the quality of life of stroke survivors [51]. An estimated 75% of survivors experience difficulties completing activities of daily living [3], which has been shown to be correlated with quality of life [52]. Therefore, improving survivors' ability to complete activities of daily living is an important rehabilitation target.

Rehabilitation treatment is limited by lack of clarity of the key ingredients to restoring daily arm use. Many upper extremity rehabilitation studies select clinical assessments of sensorimotor capacity as primary endpoints, e.g. the Fugl-Meyer Assessment and the Box and Blocks Test. These assessments grade impairment in controlled clinical and laboratory settings, and thus have been said to test the "capacity" for movement, following the taxonomy defined in the International Classification of Functioning, Disability and Health [7]. However, these scores alone do not well explain daily arm use [53], called "performance" [7]. which tends to lag motor and functional capacity [42]. In other words, assessments are unreliable proxies of real-world arm use, and by extension patient quality of life. The Threshold Hypothesis has been proposed as a mechanism for this phenomenon - that a lag in use occurs until upper extremity capability reaches a threshold at which a virtuous self-training cycle (that increases use) begins [10]. This raises the question, if a threshold of motor performance defines a training goal, how can we identify stroke survivors who can reliably meet the threshold with a given intervention, thereby attaining that goal?

Multiple systematic reviews have been conducted on arm recovery prediction [54], [55]. Coupar et al. found that baseline measures of upper limb impairment such as the NIH Stroke Scale and Fugl-Meyer Assessment, and measures of upper limb function such as the Box and Blocks Test and Nine Hole Peg Test, best predicted upper limb recovery [55]. In a separate systematic review, Chen and Winstein found that baseline measures of impairment such as deep sensation, muscle tone, strength, and active range of motion, and neurophysiological measures such as properties of a motor evoked potential (presence, amplitude, latency) and imaging outcomes best predicted recovery. Amidst the many proposed prognostic variables, it is unclear if specific mechanistic relationships dictate responsiveness to therapy. Multiple models of arm recovery have been proposed [56], perhaps most notably the proportional recovery model [57]–[59] which has been recently scrutinized statistically [60], [61], and for its inability to explain recovery for low level stroke survivors [62], potentially due to lack of corticospinal tract integrity [63]. Validity aside, these models use outcome measures of motor impairment or function, as opposed to direct measures of arm use in daily life. We interpret this shortcoming as a need to discover models that predict direct measures of real-world arm use to illuminate who will respond to therapy and inform the development of generative tools for customizing rehabilitation programs to a stroke survivors' unique abilities.

Previous modeling studies have investigated upper extremity use, or changes in use, in the subacute and chronic phases of stroke [42], [64]–[67]. Two studies quantified use with accelerometry-based measures [65], [66], one with the Motor Activity Log (MAL) self-report questionnaire [64], and two with both the MAL and accelerometry-based measures [42], [67]. With respect to therapeutic interventions, three studies only administered assessments and did not control therapy, meaning that participants each received rehabilitation services per their own personalized plans of care [42], [65], [66], and two studies administered therapeutic interventions [64], [67]. Park et al. formulated their dependent variable as attaining a binary, clinically meaningful MAL Quality of Movement subscale score of ≥ 3 after constraint induced movement therapy [64]. This raises the question how would their predictive model change if the dependent variable were instead formulated as a change in

MAL? In other words, what can we learn from participants who increased their MAL a meaningful amount, but only increased to a level less than 3 (e.g. a change from one to two)? Li et al. instead formulated their dependent variable as a binary change in Motor Activity Log ≥ 0.5 from pre to post intervention. However, 0.5 is lower than reported minimum detectable MAL changes of 0.75 to 0.84 [68], [69]. Therefore, in reviewing the literature we have not found studies that quantified meaningful changes in arm use after therapy.

The purpose of this retrospective study was to evaluate the feasibility of predicting increase in daily arm use following rehabilitation intervention with practical, widely used clinical measures, and to identify the most important measures for making such predictions with an explanatory model. We hypothesized that baseline measures of upper extremity impairment and amount of use would predict change in use following intervention. Meaning, there is an amount of use that stroke survivors with a given level of motor capacity achieve, on average. If an individual falls below that average level then they have the potential to use their arm more, which in turn could be realized through therapy. To investigate this hypothesized relationship, we merged data from several upper extremity stroke clinical trials that administered different technology-based and conventional interventions for different durations in different study environments to build an explanatory model of change in daily arm use following intervention.

Methods

We analyzed data from seven upper extremity stroke robotic rehabilitation clinical trials approved by the University of California, Irvine Internal Review Board [23], [44], [70]–[74]. All studies measured the model dependent variable (DV), the Motor Activity Log Amount of Use subscale at a baseline and a follow up timepoint, post-intervention.

The studies varied in therapy type, stroke phase, and study environment. Each evaluated a novel technology-based rehabilitation (some robotic and others sensor-based). Two of the seven clinical trials studied stroke survivors in the subacute phase of stroke (< 6 months post stroke), four studied stroke survivors in the chronic phase (≥ 6 months post stroke), and one studied both subacute and chronic phases. For the therapeutic interventions, two studies conducted them in a laboratory, one at a rehabilitation clinic, and four at participant's homes. Our analysis procedure is described below and summarized in Figure 2. All analysis was performed in MATLAB R2021a.

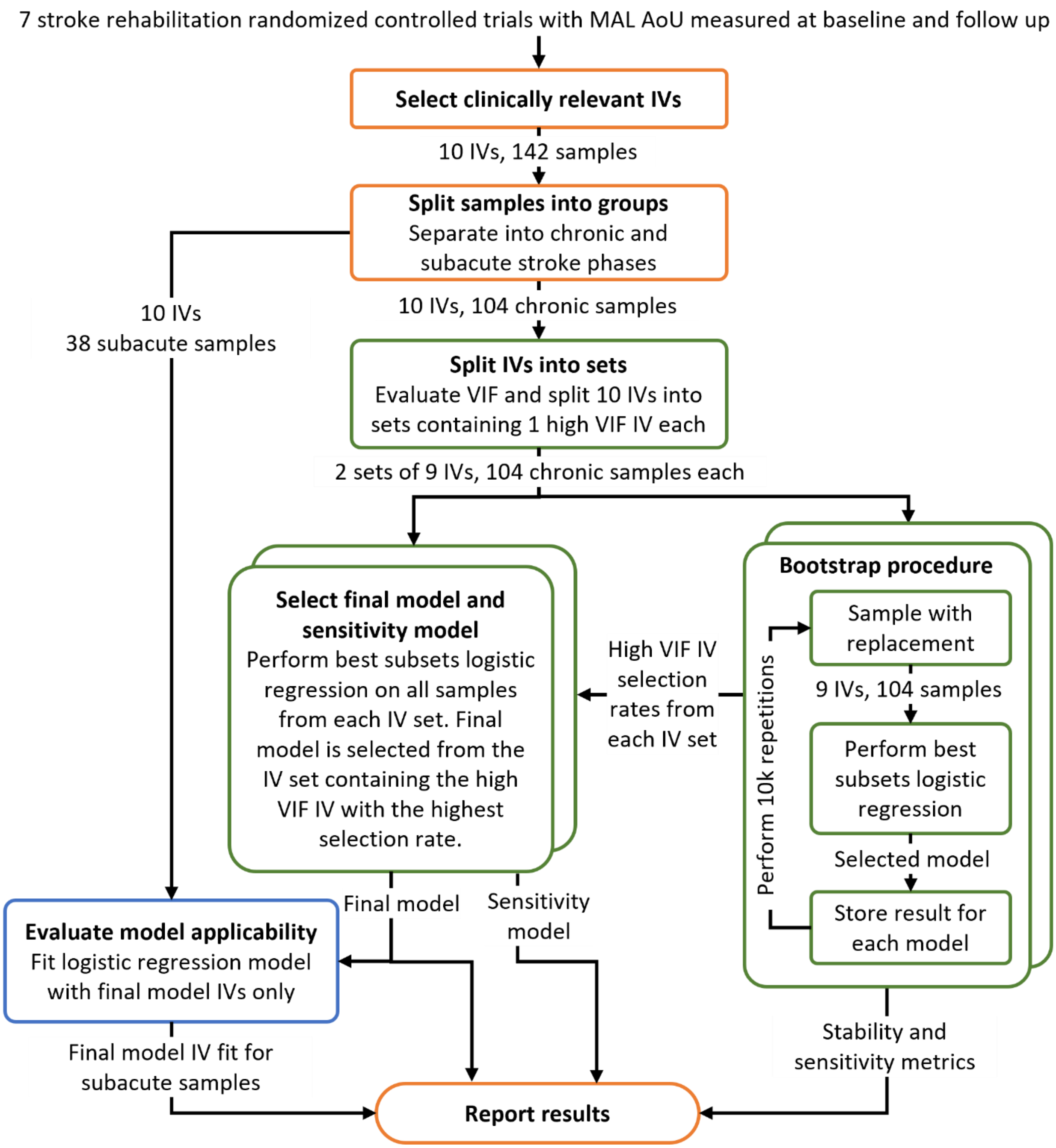

Figure 2. Analysis procedure perfromed to find an explanatory model that maps clinically relevant independent variables (IVs) measured at baseline to binary change in Motor Activity Log Amount of Use (MAL AoU).

Steps performed on both chronic and subacute stroke stage samples are outlined in orange, chronic-only in green, and subacute-only in blue. All steps were performed once, except the “Bootstrap procedure” and “Select final model and sensitivity model”, which were performed twice - once per set of reduced multicollinearity IVs.

A. Sample size

We analyzed a total of 142 participants. We assigned participants into one of two analysis groups by stroke phase, yielding 38 samples in the subacute group and 104 in the chronic group.

B. Dependent Variable

Change in arm use was quantified by the validated 30 item Motor Activity Log Amount of Use (MAL AoU) subscale [68], [75], [76], a self-rating scale quantified on a six-level scale from zero to five. We calculated a binary measure of change in MAL AoU from baseline (before therapy) to follow up (1-3 months after therapy) scores using Equation 1. For studies that collected multiple baseline timepoints, the first baseline MAL AoU was used. For crossover studies, only the baseline and follow up time points before crossover were used. The distribution of the DV is summarized in Table 1.

$$DV = \begin{cases} 1 & follow\ up - baseline\ MAL\ AoU \geq 1.0 \\ 0 & follow\ up - baseline\ MAL\ AoU < 1.0 \end{cases}$$

Equation 1. Binary measure of increasing MAL AoU from baseline to follow up.

We selected a minimum increase of 1.0 on the MAL AoU scale as indicating a meaningful increase in use. Lang et al. reported a minimal clinically important difference of 1.0-1.1 for the MAL Quality of Movement subscale in [77]. Chen et al. reported a minimal detectable change of 0.84 for the MAL AoU subscale in [69]. Similarly, Van Der Lee et al. reported that a change in MAL AoU less than 0.75 is likely to be caused by noise [68]. We thus defined a meaningful change in amount of use from baseline to follow up as a minimum increase of 1.0.

Table 1. Summary of binary dependent variable distribution, the increase in Motor Activity Log Amount of Use from baseline to follow up, per stroke phase analysis group.

Motor Activity Log Amount of Use change from Baseline to Follow up	Subacute		Chronic	
	Frequency	%	Frequency	%
< 1.0	20	52.6	77	74.0
≥ 1.0	18	47.4	27	26.0

C. Independent Variables

We selected ten clinically relevant measures [42], [64], [78]–[81] that were collected at baseline in all studies, defining these as our set of potential independent variables (IVs). We only selected variables for which data was available for ≥90% of the samples; for example stroke type (ischemic, hemorrhagic) was not used because the data were missing for several participants and would reduce the sample size. The IVs were comprised of participant demographic information (sex, age), stroke information (concordance of stroke, time after stroke), clinical measures (the Fugl-Meyer Assessment, NIH Stroke Scale, MAL AoU subscale, MAL Quality of Movement subscale, Box and Blocks Test), and the study therapeutic intervention type, as described next.

Three of the ten IVs were categorical data: sex, concordance of stroke (whether the more-affected side is the same as the dominant hand), and a latent variable of intervention type. For therapy type, we categorized each participant's treatment as conventional or technology-based. Technology based therapies used some type of novel sensor, e.g. the Music Glove [23], [44] or robotic system, e.g. the BONES [70] and FINGER [71] robots, to facilitate intense movement practice. Category frequencies for each variable are listed in Table 3.

Table 2. Summary of baseline categorical independent variables per stroke phase analysis group.

Variable	Classes	Subacute		Chronic	
		Frequency	%	Frequency	%
Sex	Female	9	23.7	32	30.8
	Male	29	76.3	72	69.2
Concordant	False	26	68.4	49	47.1
	True	12	31.6	55	52.9
Study intervention type	Conventional	16	42.1	8	7.3
	Tech	22	57.9	59	54.1

Seven of the ten IVs were continuous: age, time after most recent stroke, Fugl-Meyer Assessment on the Upper Extremity subscale, the NIH Stroke Scale, the 30 item MAL AoU, the 30 item Motor Activity Log Quality of Movement subscale, and the Box and Blocks Test (BBT) for the more stroke affected arm. We normalized BBT scores using z-scores with data from age and sex matched controls reported by Mathiowetz et al. in [82]. Means and standard deviations for each IV are listed in Table 3.

Table 3. Summary of numerical independent variables per stroke phase analysis group.

Name	Subacute		Chronic	
	Mean	SD	Mean	SD
Age [years]	55.8	12.1	58	12
Time after stroke [days]	27.9	23.8	1136	1204
Fugl-Meyer Assessment, Upper Extremity [0-66]	32.0	14.7	43	15
NIH Stroke Scale [0-42]	3.7	2.1	3	3
Motor Activity Log Amount of Use [0-5]	0.7	0.9	1.7	1.4
Motor Activity Log Quality of Movement [0-5]	0.7	0.9	1.7	1.3
Box and Blocks Test [z-score]	-6.8	2.5	-6.2	2.6

D. Model identification procedure

We found an explanatory model of the binary dependent variable in a three phased approach using data from the chronic stroke population following modeling recommendations in [83]. In the first phase we removed collinear variables using the Variance Inflation Factor (VIF) and bootstrap resampling. In the second phase, we selected a model from the reduced set of

IVs using best subsets logistic regression and evaluated its stability using bootstrap resampling. In the third phase, we evaluated the sensitivity of the model by swapping the IVs that were included and excluded in phase one and repeating the final model selection procedure in phase two. Finally, we evaluated the generalizability of the chronic phase model using the unanalyzed subacute samples.

We reduced IV multicollinearity by removing IVs. IVs with a VIF ≥ 5 were flagged as highly correlated with other IVs – we will refer to these IVs as "high VIF IVs". We then selected which high VIF IVs to remove using the following algorithmic approach. We separated high VIF IVs to generate sets with reduced multicollinearity. For example, for IVs, "A", "B", "C", and "D", if A and B had high VIF then two sets would be generated: A, C, D and B, C, D. For each set, we used best subsets regression on bootstrap resampled data to select IVs that best explained the binary DV. Unlike heuristic-based forward and backward elimination approaches, best subsets regression selects one model by comparing all possible models (combinations of IVs) - in this case $2^9 = 512$ models. Each model was selected in three steps, using the algorithm outlined in [84]. First, a logistic regression model was fit for all combinations of IVs, where the minimum model contains no IVs (a constant term only), and the maximum model contains all nine IVs. Second, models of the same size (0, 1, …, 9 IVs) were compared, and the one with minimum sum squared error of each size was advanced to the next step. Third, amongst the ten models, the model with minimum Akaike Information Criterion was selected for the current set of bootstrap resampled data. 10,000 bootstrap resamples yielded 10,000 models for each set of (reduced multicollinearity) IVs. Across the 10,000 models per IV set, we found a selection frequency for each high VIF IV and retained the set with the high VIF IV that was selected most frequently, i.e. we selected the high VIF

IV that explained the DV best, and removed the other high VIF IV that it was correlated with. Finally, we recalculated the VIF on the reduced set to confirm it had acceptable multicollinearity.

In the second analysis phase, we performed best subsets regression on the reduced set of IVs with all chronic stroke phase samples to find a final model, following recommendations in [83]. We used the 10,000 bootstrap resampling-based models from the previous phase to evaluate stability of the final model. We computed IV selection frequencies, model selection frequencies, and coefficient estimate distributions across the resampled models to evaluate final model stability.

In the third analysis phase, we evaluated the sensitivity of the final model to removing the IV in phase one. We switched the IVs that were removed and retained then repeated the second analysis phase. Note that we did not solely include the previously removed IV (and not exclude the one previously retained) to maintain acceptable multicollinearity amongst the IVs.

Lastly, to evaluate generalizability we used multiple logistic regression to test whether the IVs from the final model (selected using chronic samples only) significantly predicted the DV using the subacute stroke phase samples. We forced the time after stroke IV into the model to account for the effects of spontaneous recovery on increase in arm use. Of note, we did not perform variable selection with the subacute analysis group (as we did with the chronic) due to the small number of subacute samples (38). A popular rule of thumb is ≥ 10 samples per IV, typically termed "events per variable", to apply variable selection techniques [85], which greatly restricts the allowable number of candidate IVs for the subacute group. We therefore did not perform variable selection with the subacute samples, and only used them to assess the generalizability of the chronic model to a different (subacute) population.

RESULTS

A. Correlated independent variable removal

The MAL AoU and MAL Quality of Movement subscales had VIFs ≥ 5, with VIFs of 8.4 and 9.7, respectively. After 10,000 bootstrap resamples, the MAL Amount of Use variable was selected in 88.5% of the models, and MAL Quality of Movement in 80.1%. Therefore, MAL Quality of Movement was removed. Removal reduced the VIF of MAL Amount of Use to an acceptable magnitude of 2.6. The VIF for each variable is listed in Table 4.

Table 4. Variance inflation factors before and after removing Motor Activity Log Quality of Movement to reduce multicollinearity in the chronic stroke phase analysis group.

Name	Variance Inflation Factor	
	Initial	After removing MAL QoM
Sex	1.4	1.4
Concordant	1.2	1.1
Intervention type	1.1	1.1
Age in years	1.3	1.3
Time after stroke	1.1	1.1
Fugl-Meyer Assessment, Upper Extremity	3.4	3.4
NIH Stroke Scale	1.5	1.5
Motor Activity Log Amount of Use	8.4	2.4
Motor Activity Log Quality of Movement (MAL QoM)	9.7	
Box and Blocks Test [z-score]	3.4	3.3

B. Chronic phase model selection

A candidate set of nine IVs (11.6 events per variable with MAL Quality of Movement removed) was used for variable selection. Best subsets logistic regression selected a final model with two IVs: the BBT z-score and MAL AoU subscale at baseline (Table 5).

Holding baseline MAL AoU constant, a one unit increase in BBT z-score raised the odds of increasing MAL AoU by 161% (95% CI [66%, 311%]). Holding baseline BBT constant, a one unit increase in baseline MAL AoU reduced the odds of increasing MAL AoU by 50% (95% CI [9%, 73%]). Model estimated probabilities of increasing MAL AoU and the raw DV data (increased or did not increase MAL AoU) are plotted against the two IVs in Figure 3.

Table 5. Coefficient estimates for the global and final selected models.

Model variable	Global model			Final model			Bootstrap derived coefficient estimates		
	Coeff	SE	p	Coeff	SE	p	Median	2.5^{th}	97.5^{th}
Intercept	9.26	4.00	0.02	5.42	1.68	0.00	8.57	-1.93	23.39
MAL AoU	-0.90	0.36	0.01	-0.70	0.31	0.02	-1.04	-2.71	-0.50
BBT	1.13	0.29	0.00	0.96	0.23	0.00	1.24	0.58	2.52
Sex	-1.18	0.74	0.11				-1.90	-3.80	-0.99
Concordant	0.45	0.65	0.48				1.33	-1.91	3.81
Intervention	-0.14	0.62	0.82				-1.02	-2.89	2.26
Age	-0.05	0.03	0.11				-0.08	-0.19	-0.04
Time after	0.00	0.00	0.39				0.00	0.00	0.00
FMAUE	0.02	0.03	0.51				0.07	-0.10	0.20
NIHSS	0.14	0.14	0.33				0.30	0.15	0.69

Abbreviations: Coeff, coefficient estimate; SE, standard error; p, p-value; 2.5^{th}, 2.5^{th} percentile of the coefficient estimates over 10,000 bootstrap resamples; 97.5^{th}, 97.5^{th} percentile of the coefficient estimates over 10,000 bootstrap resamples; MAL AoU, Motor Activity Log Amount of Use subscale; BBT, Box and Blocks Test z-score; Intervention, classification of intervention type received during the study: conventional or technology-based; Time after, time after stroke; FMAUE, Fugl-Meyer Assessment, Upper Extremity; NIHSS, National Institute of Health Stroke Scale.

Applying decision threshold of 0.5 to model estimated DVs yielded a classification accuracy of 74%, with 34% sensitivity and 89% specificity (a model output ≥ 0.5 corresponded to a participant increasing MAL AoU after therapy). Note that the disparity between sensitivity and specificity is in part due to the imbalance in the DV class labels (26% positive class, 74% negative), as logistic regression equally weights all samples in its cost function.

Across all bootstrap resampled models, the final model was selected most frequently (8.6% of resamples). The next most frequently selected model incorporated two additional IVs, the Fugl-Meyer Assessment and NIH Stroke Scale, and was selected approximately half as often (4.4%). Coefficients and standard errors for the global model (all nine IVs) and the final model are listed in Table 5 with bootstrap derived coefficient estimates. Resampling selection rates for each IV and the top ten models across the 10,000 bootstrap resamples are listed in Table 6.

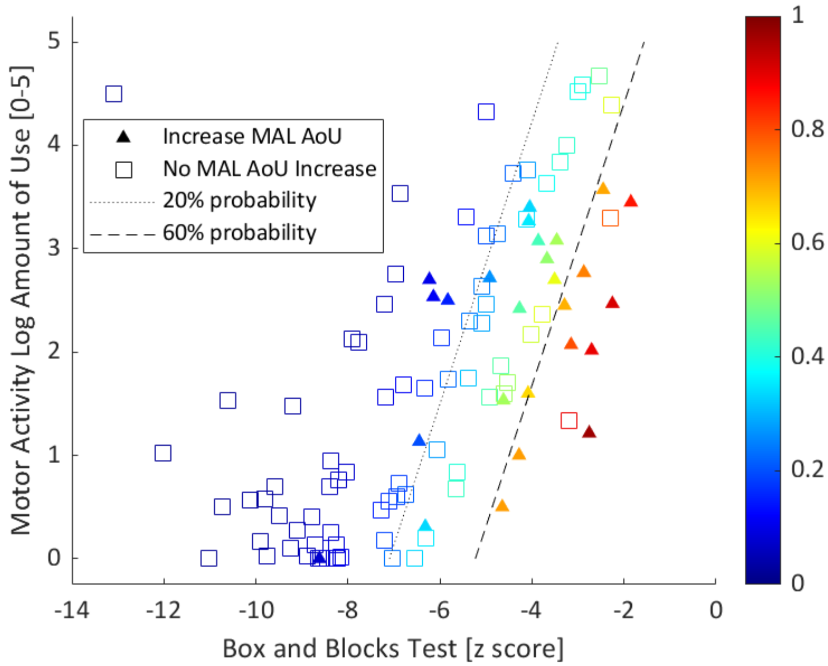

Figure 3. Model estimated probabilities of increasing Motor Activity Log Amount of Use after intervention and raw data plotted against model independent variables for chronic samples. Four data dimensions are displayed. The x and y axes, baseline Box and Blocks Test (BBT) z-score and baseline Motor Activity Log Amount of Use (MAL AoU) subscale, are the independent variables of the final model. The model estimated probabilities of the dependent variable, increasing MAL AoU after intervention, are plotted in color temperature. Hotter colors correspond to higher model estimated probabilities of increasing MAL AoU after intervention and colder colors correspond to higher estimated probabilities of not increasing MAL AoU after intervention. The dashed and dotted lines are constant model estimated probabilities, e.g. the model estimated probability of increasing MAL AoU is 60% along the dashed line. The raw binary dependent variable values are encoded by marker shape. Participants who did increase MAL AoU after intervention are plotted as solid triangles, and participants who did not are plotted as unfilled squares. Comparing color temperature and marker shape, the model estimates whether participants increased MAL AoU well for participants with mismatched baseline BBT and MAL AoU – most participants to the right of the 60% constant probability line increased MAL AoU after intervention, had high baseline BBT, and relatively low baseline MAL AoU.

Table 6. Selection rates for each independent variable and the top ten selected models after 10,000 bootstrap resamples.

		Interv.	Time	Conc.	FMAUE	NIHSS	Age	Sex	**MAL AoU**	**BBT**
		Independent variable selection rate [%]								
		21.1	28.3	34.1	36.5	39.9	46.9	47.9	**88.5**	**99.9**
	Model selection rate [%]	Independent variable in model?								
1	8.6								**X**	**X**
2	4.4				X	X			**X**	**X**
3	3.7						X	X	**X**	**X**
4	3.6				X				**X**	**X**
5	3.2			X			X		**X**	**X**
6	2.9						X		**X**	**X**
7	2.7			X			X	X	**X**	**X**
8	2.5					X			**X**	**X**
9	2.1							X	**X**	**X**
10	2.1		X				X	X	**X**	**X**

Independent variable selection rates across all models are listed at the top of the table, and the top ten selected models are described in the remaining portion. For each model, cells marked with an "X" indicate that the independent variable was included in the model by the best subsets logistic regression method. For example, the model selected most often contained the two IVs MAL AoU and BBT and was selected in 8.6% of the 10,000 bootstrap resampled models. The IVs in the final model (MAL AoU and BBT) are bolded and were selected in all top ten models. Abbreviations: Interv., classification of intervention type received during the study: conventional or technology-based; time, time after stroke; Conc, concordance of stroke; FMAUE, Fugl-Meyer Assessment, Upper Extremity; NIHSS, National Institute of Health Stroke Scale; MAL AoU, Motor Activity Log Amount of Use subscale; BBT, Box and Blocks Test z-score.

C. *Final model sensitivity to correlated independent variable removal*

We switched which IV was removed (MAL Quality of Movement) and retained (MAL AoU) and repeated both the model selection procedure and stability analysis. Best subsets logistic regression again selected a model with two IVs: the BBT z-score (same as final model) and MAL Quality of Movement subscale at baseline. Across all bootstrap resampled models, the final model was again selected most frequently (8.8% of resamples). Coefficients and standard errors for the global model (all nine IVs) and the final model are listed in Table 7, with bootstrap derived coefficient estimates. Resampling selection rates for each IV and the top ten models across the 10,000 bootstrap resamples are listed in Table 8.

Table 7. Sensitivity analysis coefficient estimates for the global and final selected models.

Model variable	Global model			Final (sensitivity) model			Bootstrap derived coefficient estimates		
	Coeff	SE	p	Coeff	SE	p	Median	2.5th	97.5th
Intercept	8.77	3.95	0.03	4.99	1.64	0.00	7.76	-1.26	21.74
MAL QoM	-0.74	0.35	0.03	-0.62	0.30	0.04	-0.92	-2.10	-0.46
BBT	1.06	0.28	0.00	0.91	0.23	0.00	1.14	0.57	2.25
Sex	-1.16	0.72	0.11				-1.77	-4.03	-0.95
Concordant	0.28	0.62	0.65				1.20	-2.45	3.60
Intervention	-0.19	0.61	0.65				-1.04	-3.07	2.15
Age	-0.05	0.03	0.12				-0.07	-0.17	-0.04
Time after	0.00	0.00	0.36				0.00	0.00	0.00
FMAUE	0.02	0.03	0.58				0.07	-0.10	0.24
NIHSS	0.11	0.14	0.45				0.29	0.17	0.60

Abbreviations: Coeff, coefficient estimate; SE, standard error; p, p-value; 2.5th, 2.5th percentile of the 10,000 bootstrap resampled coefficient estimates; 97.5th, 97.5th percentile of the 10,000 bootstrap resampled coefficient estimates; MAL QoM, Motor Activity Log Quality of Movement subscale; BBT, Box and Blocks Test z-score; Intervention, classification of intervention type received during the study: conventional or technology-based; Time after, time after stroke; FMAUE, Fugl-Meyer Assessment, Upper Extremity; NIHSS, National Institute of Health Stroke Scale.

Table 8. Sensitivity analysis selection rates for each independent variable and the top ten selected models after 10,000 bootstrap resamples.

		Interv.	Time	Conc.	FMAUE	NIHSS	Age	Sex	MAL QoM	BBT
		Independent variable selection rate [%]								
		21.1	30.2	26.0	30.2	27.2	46.3	52.5	80.1	99.7
	Model selection rate [%]	Independent variable in model?								
1	8.8								X	X
2	6.2						X	X	X	X
3	3.6			X			X	X	X	X
4	3.5				X				X	X
5	3.3		X				X	X	X	X
6	2.7							X	X	X
7	2.1			X			X		X	X
8	1.9									X
9	1.9						X		X	X
10	1.9					X		X	X	X

Independent variable selection rates across all models are listed at the top of the table, and the top ten selected models are described in the remaining portion. For each model, cells marked with an “X” indicate that the independent variable was included in the model by the best subsets logistic regression method. For example, the model selected most often contained the two IVs MAL QoM and BBT and was selected in 8.8% of the 10,000 bootstrap resampled models. Abbreviations: Interv., classification of intervention type received during the study: conventional or technology-based; time, time after stroke; Conc, concordance of stroke; FMAUE, Fugl-Meyer

Assessment, Upper Extremity; NIHSS, National; MAL QoM, Motor Activity Log Quality of Movement subscale; BBT, Box and Blocks Test z-score.

D. Generalizability of the model to a subacute sample

We used logistic regression to analyze the relationship between the same DV, binary increase in MAL AoU after intervention, the IVs from the final chronic stroke phase model, BBT and MAL AoU baseline, and the time after stroke IV, for the subacute stroke phase samples. All p-values indicate a non-significant relationship, with p=0.76 for MAL AoU and p=0.21 for BBT, although time after stroke had a near significant relationship (p = 0.06). The model coefficient estimates, standard errors, and p-values are listed in Table 9.

Table 9. Final logistic regression model fit using subacute stroke stage samples.

Independent variable	Coefficient estimate	Standard error	p-value
Intercept	1.28	1.19	0.28
Motor Activity Log Amount of Use	-0.12	0.39	0.76
Box and Blocks Test [z-score]	0.19	0.15	0.21
Time after stroke	-0.04	0.02	0.06

The independent variables were selected using all the samples from the chronic stroke stage participants. The model was fit using all the samples from the subacute stroke stage.

DISCUSSION

We developed a model that estimates whether the MAL AoU score will increase after intervention using data from seven upper extremity, technology-based, stroke rehabilitation clinical trials. The studies varied in intervention type (conventional and different types of novel technology-based interventions), intervention dose, study environment, and participant stroke phase. We selected a model in three phases using all chronic stroke phase samples first, then evaluated its generalizability to the unanalyzed subacute samples. In the first phase, we identified IV multicollinearity and removed correlated variables using a combination of the VIF and bootstrap resampling. The MAL AoU and Quality of Movement subscales were found to be correlated and the Quality of Movement IV was removed using an algorithmic, bootstrap resampling-based procedure. Second, we performed IV selection using best subsets logistic regression to find a final model comprised of two IVs, MAL AoU and BBT z-score. We evaluated final model stability using results from the bootstrap resampling performed during phase one and found that final model and its IVs were selected most often. Third, we evaluated model sensitivity by repeating the phase two model selection procedure using the previously excluded MAL Quality of Movement instead MAL AoU. A similar model was selected with equivalent IVs, similar coefficient estimates, and similar stability (similar final model and IV selection rates from bootstrap resampling). Lastly, using only the IVs from the final model and the unanalyzed subacute samples, we fit a new logistic regression model and found that the IVs did not significantly predict the presence or absence of increase in MAL AoU. We next discuss these results as well as limitations and future directions.

A. *The importance of "untapped use potential"*

In the final model, baseline BBT was found to be positively correlated with increasing arm use after intervention, whereas MAL AoU was negatively correlated. This model highlights a potential driver of increasing arm use: those with a mismatch in MAL AoU and BBT (specifically, unusually low MAL AoU relative to BBT z-score) tended to increase MAL AoU. We find this result intuitive, and it supports our hypothesis - those with "untapped use potential" tended to increase daily use after therapeutic intervention. This relationship can be visualized in Figure 3, where all but two participants with ≥ 60% model estimated probability of increasing MAL AoU (samples falling on and to the right of the dashed line) did in fact increase MAL AoU (triangles). This mismatch in dexterity and use could arise from compensatory strategies, and an eventual pattern of learned nonuse or bad-habit [53], [86]. This relationship did not transfer to the subacute sample for which increase in use is more likely to be explained by time after stroke, a negative and near significant ($p = 0.06$) IV in the subacute model. For chronic stroke survivors that did benefit from the intervention, intervention type (a binary IV with labels conventional and technology-based intervention) was not a statistically significant predictor ($p = 0.82$), and it was selected least amongst all candidate IVs during bootstrap resampling (included in 21.1% of models), suggesting that individuals with untapped use potential may benefit from any reasonable intervention, conventional or technology-based. This however does not imply the converse, that individuals without untapped use potential (matched or high MAL AoU relative to BBT) will not benefit from any type of intervention. For these individuals, the model suggests that rehabilitation should target increasing BBT, where a one unit increase in BBT z-score raised the odds of increasing MAL AoU by 161%, which has been shown to benefit from technology-based high intensity training [44], [71], [87]–[89]. Taking these ideas together, stroke survivors with untapped use potential could potentially benefit from simple, low-tech therapies that shape behavior, while those without untapped use potential could benefit from therapies that target arm dexterity using interventions tailored to their impairment.

B. Final model sensitivity

The final model was sensitive to which IV was removed to reduce multicollinearity. Removing baseline MAL Quality of Movement resulted in a final model containing baseline BBT z-score and MAL AoU, whereas instead removing baseline MAL AoU unsurprisingly resulted in the opposite – the final model contained baseline BBT z-score and MAL Quality of Movement. Although the MAL has been reported to be both reliable and valid [68], [76] high collinearity between the two subscales has been reported [76] consistent with the findings in this study. The coefficient estimates for each final model were similar: 0.96, -0.70

for BBT and MAL AoU, respectively and 0.91, -0.62 for the BBT and MAL Quality of Movement sensitivity analysis IVs, respectively. Further, the stability analyses for each set of IVs (each including one of the MAL subscales) yielded similar results: both the MAL subscales were included in >80% of the models, second only to BBT z-score (included in ≥99.7% of the models), and in both cases the final model was selected most frequently (8.6% and 8.8% model selection rates for the MAL AoU IV set and MAL Quality of Movement IV set, respectively). Considering that the two subscales appear to measure non-orthogonal dimensions and that selected models were both similar and stable regardless of which MAL subscale was used, the results indicate that some combination of upper extremity movement quality and amount, in addition to dexterity, are important for explaining a stroke survivor's potential for increasing use after therapy.

C. *Use of Motor Activity Log*

Baseline MAL AoU was used to calculate the binary DV, change in MAL AoU from baseline to follow u, and was algorithmically included in the final model as an IV, which could be a result of the statistical phenomenon commonly known as "regression toward the mean". There is a potential that regression toward the mean influenced the model in two ways. First, MAL Quality of Movement was algorithmically removed (instead of MAL AoU) from the candidate set of IVs to reduce IV multicollinearity. However, doing the opposite and instead excluding MAL AoU would not ultimately alleviate the potential effects of regression toward the mean due to collinearity between the two subscales. Indeed, when instead excluding MAL AoU in the sensitivity analysis, MAL Quality of Movement was selected in the final model (if MAL Quality of Movement were not selected then it would not be falsely included in the model due to regression toward the mean, at least for the Quality of Movement subscale). The second potential influence of regression toward the mean was during final model selection using the best subsets logistic regression method – models with MAL AoU or Quality of Movement could have been selected over others because their correlation with the DV is falsely inflated. We attempted to limit the risk of falsely including these IVs in logistic regression models by computing the DV with a change threshold of ≥ 1.0 based on reported MAL variability [68], [69], [77]. A more robust and modern approach would be to replace the subjective MAL with an objective measure of activity. Estimation of clinically meaningful activity-based metrics from wearable device trackers, like a wrist worn accelerometer, is an open area of research and a challenging problem [18], [78], [90]–[93]. However, given such metrics, a stable baseline of activity could theoretically be established per participant and a statistically significant threshold of activity change more accurately estimated. Further, if movement quality and amount are indeed independent measures, then orthogonal ones could be more easily developed using such objective measurement-based metrics, also addressing the model sensitivity shortcomings we previously discussed. Considering the limitations of the MAL as measurement tool, in important next step in investigating the interplay between upper extremity dexterity and activity is to replace the MAL with a more reliable and valid objective measure.

D. Predictors of increase in upper extremity use

The MAL AoU and BBT were algorithmically selected as statistically significant predictors of binary increase in MAL AoU from a limited set of candidate IVs. Predictors of upper extremity performance not in our candidate IV set include grip strength, participation (a person's engagement in meaningful life situations) quantified by the participation domain of the Stroke Impact Scale, the Wolf Motor Function Test, muscle tone quantified by the Ashworth

Scale, functional ability quantified by the Functional Independence measure, tactile sensation, and proprioception [42], [64], [80], [94], [95]. Some of these predictors were measured at baseline in only some of the studies included in this analysis, such as grip strength, muscle tone, and proprioception. Including these measures would have decreased our effective sample size and therefore the number of allowable candidate IVs due to limitations on the minimum number of samples/events per variable. The final model IVs BBT and MAL AoU are therefore by no means an exclusive list of variables that explain change in upper extremity amount of use. However, it is significant that a powerful explanatory model can be built with just these two, easy-to-obtain, widely used clinical scores.

This model aligns with previous studies that have developed models to identify predictors of arm use performance [42], [64]–[67]. Each found that measures of arm capacity and impairment measured at baseline most significantly predicted use. Rand et al. modeled changes in use measured by the MAL using multiple clinical assessments following hospital discharge and 12 months after stroke. They found that age, grip strength, and the Action Research Arm Test measured at discharge were the most significant predictors of MAL score [42]. In a recent study, Lang et al. modeled trajectories of impairment, capacity, and performance in a longitudinal study of subacute stroke survivors over 24 weeks and found that plateaus in performance, measured by the Action Research Arm Test, preceded plateaus in capacity, measured by an accelerometry-based arm use ratio [65]. Also in a recent study, Lundquist et al. found that the Fugl-Meyer Assessment measured at baseline was the most significant predictor of use, quantified by an accelerometry-based measures [66]. In line with our results, capacity did not guarantee performance, meaning that some participants had mismatched, greater capacity than use, whom our model identifies as having the highest probability of increasing use after therapy. Similar to the outcome measure used in our model, Li et al. found that Fugl-Meyer Assessment scores most significantly predicted change in MAL AoU, where change was also formulated as a binary measure, only with a less restrictive classification boundary of MAL AoU $\geq$0.5 in a sample of 94 chronic stroke survivors [67]. Also using the MAL as an outcome, Park et al. modeled whether participants attained a clinically meaningful MAL Quality of Movement subscale score of ≥ 3, two weeks and twelve months following constraint induced movement therapy [64]. In line with other models, the Fugl-Meyer Assessment was a positive predictor. However, contrary to our own model, MAL Quality of Movement was a positive predictor: a one unit increase in baseline MAL Quality of Movement led to a 5.2 times higher probability of attaining an MAL ≥ 3, whereas in our model MAL is a negative predictor of MAL change. This could be attributed to differences in the formulation of the DV, where our model requires a change of at least one, and the logistic regression model developed by Park et al. requires attaining an “absolute” score. Together, these studies indicate that capacity is an important predictor of performance and tends to lead changes in use.

Conclusion

We found that baseline measures of upper extremity dexterity and daily use explained whether chronic stroke survivors increased daily use after intervention, robotic or conventional. Specifically, participants with high dexterity relative to daily use tended to improve independently of the type of study intervention, which we interpret as an “untapped use potential”. This relationship, however, did not hold in a small sample of subacute participants, perhaps the most clinically relevant population, supporting the idea that untapped use potential may be a consequence of learned nonuse. Still, identifying stroke survivors with untapped use potential is a tractable problem given the prevalence of wearable devices, and could be a practical application of rehabilitation activity tracking research. Further, the identified model highlights the potential for an assessment-based approach to identify rehabilitation targets – a means of tailoring therapy to a stroke survivor's unique abilities.

CHAPTER 3: A NOVEL ROBOTIC APPROACH TO PROPRIOCEPTION TRAINING

SUMMARY OF THE CHAPTER

Proprioceptive deficits are common after a stroke and are thought to negatively impact motor learning. Despite this, there is a lack of practical robotic devices for assessing proprioception, as well as few robotic rehabilitation techniques that intensely and engagingly target proprioception. This chapter first presents a novel proprioceptive gaming paradigm that we call "Propriopixels" and the first Propriopixels game Proprioceptive-Pong (P-Pong). In P-Pong, players must continuously make game decisions based on sensed index and middle finger positions, as the game robotically moves their fingers instead of screen pixels to express the motion of the ball and paddle. We then present the design pattern of a binary impedance robot, PINKIE, developed to assess finger proprioception and play Propriopixels games. PINKIE uses low-cost actuators and sensors and is fabricated completely from 3D printed, laser cut, and off-the-shelf components. We also report the results of a pilot study in which we investigated the effect of a short bout of P-Pong play on proprioceptive acuity, and quantified user engagement and intrinsic motivation of game play. We randomly assigned 15 unimpaired human participants to play 15 minutes of P-Pong (proprioceptive training group) or a similar but video-only version of Pong (control group). We assessed finger proprioception acuity before and after game play using the Crisscross assessment previously developed by our laboratory, engagement using the User Engagement Scale, and motivation using the Intrinsic Motivation Inventory survey. Following game play, there was a significant improvement in proprioceptive acuity (2.2 ± 2.6 SD mm, $p = 0.023$) in the proprioceptive training group but not the control group (0.5 ± 0.9 SD mm, $p = 0.101$). Participants rated P-Pong highly on most survey subscales, and as highly

as visual Pong, except in the Perceived Usability and Competence subscales, a finding we discuss. To our knowledge, this work presents the first computer gaming approach for providing intense and engaging finger proprioception training, by splitting the feedback of game elements between the visual and proprioceptive senses. The pilot experiment indicates that the human sensory motor system has the ability to at least temporarily improve proprioception acuity with such game-based training.

INTRODUCTION

Upper limb motor and sensory deficits are common in stroke survivors [96], [97] and limit the ability to perform activities of daily living [98]. Proprioception has been identified as an essential input for learning [99] and a strong behavioral predictor for motor gains in the hand following constrain-induced therapy [100] and robotic hand therapy in chronic stroke [41], [71], suggesting that the training and improvement of proprioception could improve motor learning and recovery [101]. However, investigating the role of proprioception in motor learning is limited by a lack of reliable, sensitive assessments to classify and grade proprioceptive deficits [102], and by the lack of training interventions to reduce those deficits [36].

Considering the large number of practice repetitions required for sensory [13] and motor [11], [12] learning, the current model of clinical care may limit functional recovery since studies show few repetitions are practiced during therapy sessions [14]. Further, due to the relatively short patient-clinician therapy durations [14], home-based therapy is presently a major aspect of stroke rehabilitation. And yet, the adherence rate for home-based rehabilitation is low [15], potentially due to reduced motivation [103], [104]. Home-based therapy outcomes may be ameliorated by game-based therapy, which has been shown to increase motivation and repetitions [43], [44]. We interpret these shortcomings as indicating a need for practical, patient-accessible, motivating, proprioception training interventions, and developed the Phalange traINer for KInesthesia and Extension (PINKIE) as well as a proprioceptive gaming paradigm, that we call Propriopixels, as a potential solution.

PINKIE is a simplified, compact version of the FINGER robot [105]. We implemented the Crisscross proprioception assessment, previously developed for FINGER [106], on PINKIE and created Proprioceptive Pong (P-Pong), a computer game that specifically targets proprioceptive acuity training using the Propriopixels paradigm. While games have been developed that incorporate simultaneous sensory feedback (vision, touch, proprioception) of game elements [38]–[40], previous approaches do not integrate proprioception sensing as a required input to game decisions the player makes to succeed. As we describe next, we designed P-Pong to bring the benefits of gamification to finger proprioception training, with the goal of provoking somatosensory learning with motivating, high-intensity training that is continuously and explicitly focused on finger position sensing in order to successfully play the game.

METHODS

A. The Propriopixels Gaming Paradigm and Proprioceptive Pong Development

We developed Propriopixels as a general-purpose method for targeting proprioception during game-based training by requiring the use of proprioceptive afferents to make game play decisions. In a typical video game, all game elements are shown on a screen, so the player uses vision to sense how elements evolve and make corresponding decisions. In the Propriopixels paradigm, one of the game dimensions is conveyed to the finger with a robotic device instead of the screen, therefore requiring the use of proprioception to play. That is, we create a "Propriopixel" by moving the finger instead of a light pixel on screen. We implemented this paradigm in P-Pong.

P-Pong is based on the popular Atari arcade game. To make it a Propriopixels game, feedback of the player paddle and ball are divided between a screen and robot manipulandum (Figure 4). The manipulandum drives the player's middle finger according to the ball's position in the virtual field. The player then moves the index finger, trying to match the position of the index "paddle" finger with that of the middle "ball" finger to hit the ball. From a traditional psychometric perspective, the game is akin to a joint position reproduction task [107]: overlapping the index with the middle finger in physical space corresponds to aligning the paddle with the ball in virtual game space. Differing from a typical joint position reproduction task, the middle finger target position is dynamic, constantly presented (requiring no memorization from the player), and reproduced with an ipsilateral finger [107].

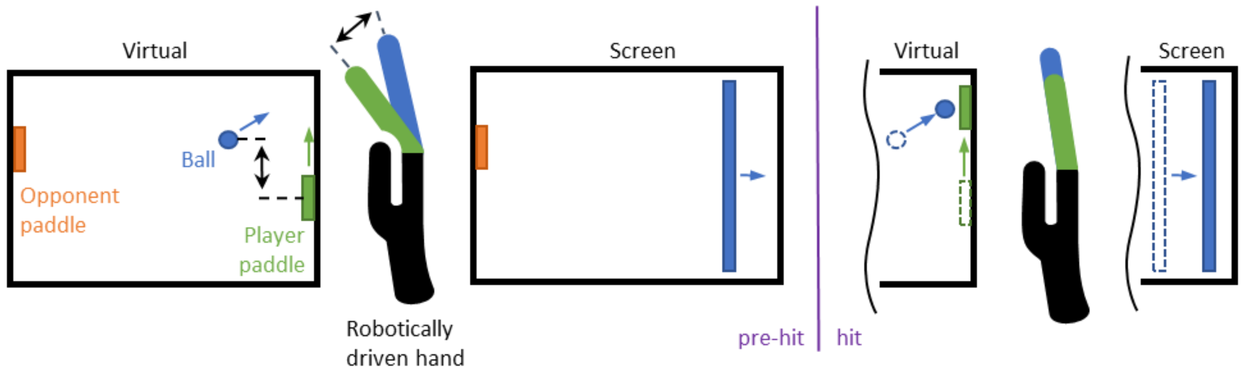

Figure 4. Propriopixels and Pong: how the display of the ball and paddle are split between the robot and screen.

Snapshots of game play are shown at two timepoints “pre-hit” (left) and “hit” (right). At pre-hit (left) the ball is travelling toward the player’s paddle, shown in the virtual game state, where by “virtual” we mean the internal game state that is not displayed on the screen but is instantiated in code. The player senses the positions of game elements through separate modalities, either proprioceptively from their robotically-driven hand or visually from the screen. The robotically-driven hand conveys the vertical position of the ball by moving the middle finger and vertical position of the paddle by moving the index finger. The screen conveys the vertical position of the opponent paddle and the horizontal position of the ball as vertical lines. To successfully hit the ball, the player must drive the paddle finger to overlap with the ball finger (pre-hit to hit timepoints). The overlap generates a hit (right) and returns the ball. The paddle, vertical ball position, and player’s robotically-driven hand are visually hidden while the ball travels toward the player, and all reappear when the ball reverses. The paddle can be controlled via active movements of the index “paddle” finger which backdrive the robot, or with the contralateral thumb and joystick using the handheld controller, which controls the robot paddle finger motor. The clutch is disconnected for the former and coupled to the finger mechanism for the latter passive “paddle” finger movement scheme. The joystick input mode is intended to make gameplay accessible to people with a stroke who cannot actively move their fingers and was the mode evaluated in this study.

We replaced the classic ping pong match-style play with a survival style match. At the start of the match the player paddle is wide, and the ball speed is slow. As the player “survives” by continuing to return the ball, the ball speed increases, and the paddle width decreases, until reaching preset limits. The game-controlled opponent always returns the ball, and the game ends when the player misses the ball. From a proprioception perspective, the game difficulty ramps up from start: decreasing the paddle width corresponds to decreasing the allowable separation distance, or position error, between the fingers. We found this mode more engaging than traditional match play and adaptation has been reported to improve reinforcement learning [108].

B. Development of the Binary Impedance Robot PINKIE

Propriopixels gives rise to a simple robot design pattern, which we call a "binary impedance robot". Compared to more common assistive and rehabilitative robot designs that utilize high-cost actuators and sensors to render a continuum of impedances, a binary impedance robot only renders the two limits – high and low impedance. It is either stiff to passively drive the Propriopixel finger, or transparent to sense active finger movements for a game input.

We developed the binary impedance robot PINKIE as a practical, low-cost device for the assessment and training of finger extension and proprioception. Like a typical video gaming system, it has a console (PJRC *Teensy 3.6*), display (Excamera Labs *Gameduino 3X*), and handheld controller (Nintendo *Wii Nunchuk*). Proprioception is engaged through the manipulandum (Figure 5) which can act as a user input device, sensing active finger movement, or output, actuating the fingers in unfurling/furling trajectories. With such a system we can substitute visual states for proprioceptive ones, i.e. we move fingers instead of screen pixels to express the motion of game elements.

The index and middle fingers are each guided by a prismatic-revolute mechanism. For passive finger movements, the linear degree of freedom is coupled to a lead screw actuator (Actuonix Motion Devices *P16*) through a simple magnetic clutching system (Figure 6). For low-resistance active movements the position dependent clutch is disconnected: the actuator retracts past the range of motion of the mechanism, thereby pulling the mechanism against its flexion hard stop and separating the magnets affixed to the mechanism and actuator, leaving the mechanism free for patient driven movements.

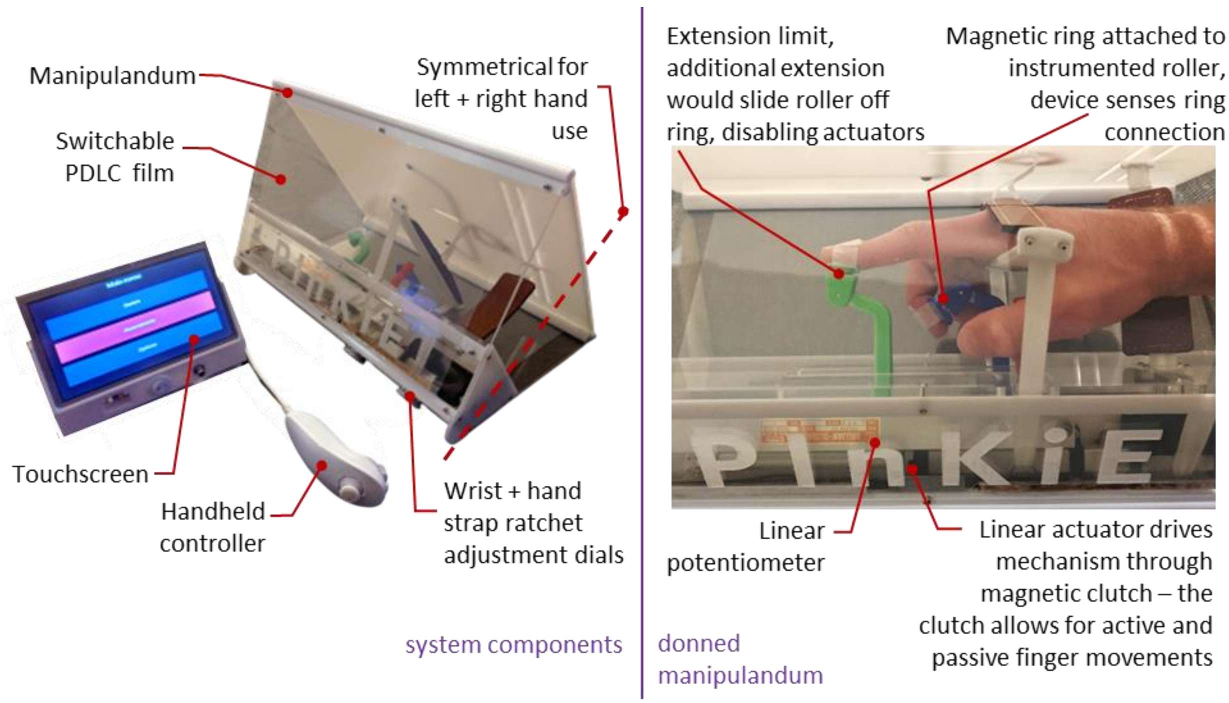

Figure 5. PINKIE system description.

Layout of PINKIE system components (left) and closeup of the donned manipulandum (right). In both images the PDLC film that covers the top and bottom of the manipulandum is transparent. All device functionality such as user settings, assessments, and training games are controlled via the touchscreen.

Donning involves rotating the entire device about its long axis to the left- or right-hand orientation, taping magnetic rings to the fingers, and adjusting support straps at the hand and wrist with ratchet dials (BOA *S2*). The magnetic finger “rollers” (Figure 5) attract the rings to simplify donning, act as mechanical fuses to protect against over-extending the fingers, and sense ring presence to enable the actuators through an embedded switch. The device is covered with electrically switchable PDLC film (Smart Tint *LV-SF*) to control hand visibility with no moving parts; the microcontroller can instantaneously switch the film from transparent to opaque. Of note, we fabricated all custom components solely using 3D printing and laser cutting.

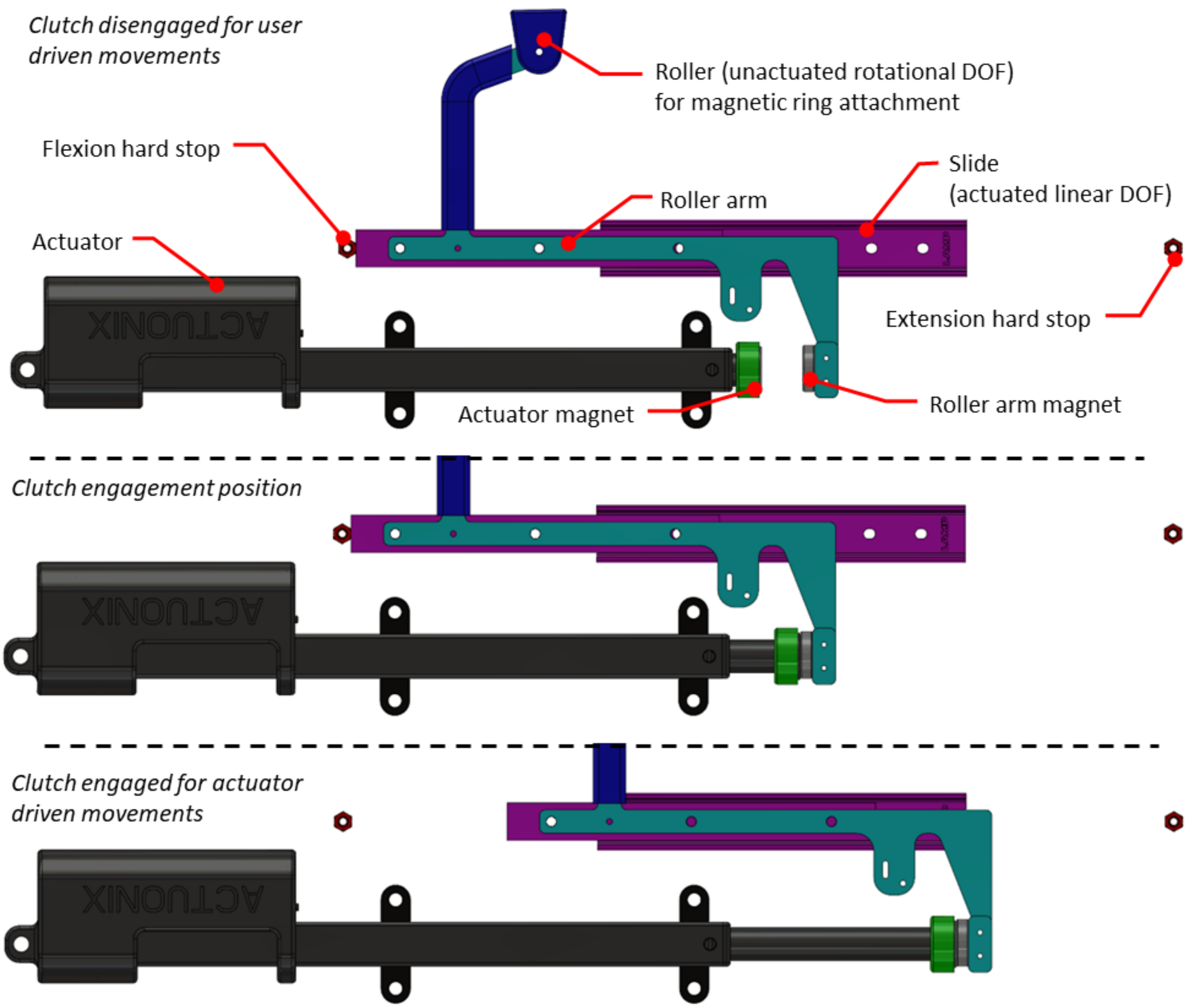

Figure 6. Magnetic position clutch functionality.

The simple magnetic clutch makes PINKIE a binary impedance robot. When the clutch is disengaged (top), the user can move the mechanism between the two hard stops. The only resistance to motion is due to the small amount of mechanism inertia and friction in the slide and roller. In the shown position, the user's finger is fully flexed and the force from the actuator magnet is undetectable. Linear roller arm movement is sensed using a linear potentiometer (P3 America CFL Series) located out of plane. To engage the clutch (middle), the actuator rod is extended toward the roller arm magnet. The clutch can be engaged in any position throughout the range of motion of the mechanism simply by moving the actuator close to the roller arm magnet. When engaged (bottom), the actuator drives the linear degree of freedom. In the finger extension direction, the roller arm is driven via normal force between the two magnets, and in the flexion direction it is driven via magnetic force. This force configuration is especially advantageous for moving the fingers of stroke survivors who commonly have more resistance to finger extension than flexion due to high flexor muscle tone.

We developed a control algorithm to track velocity trajectories used for robotic rehabilitation activites such as the ball Propriopixel finger movement in P-Pong and proprioception assessments. To achieve good tracking with low cost actuators and drive electronics, we found a nonlinear, feedforward model of the plant to compensate for the nonlinearities that the plant exhibits - deadband and hysterisis due to the lead screw based design of the actuator.

We performed system identification to find a model that maps a desired state to a pulse width modulation (PWM) control command. The PWM does not encode a position as in many hobby-grade servo motors, it simply modulates the duty cycle of a constant voltage across a permanent magnet direct current motor, thus modulating the net current through the motor windings. With the robot donned and hand relaxed, the plant was excited with a pseudorandom non-binary signal ranging from 0 to 100% PWM. This input was applied to both finger mechanisms, and their linear positions were measured at 300 Hz using the mechanism linear potentiometers. The input generated a wide range of velocities (-2.5 to 2.5 in/s) and accelerations (-40 to 40 in/s^2). Raw input-output data is shown in Figure 7. To select a model, we built a library of candidate models, fit the raw input-output data to each, and selected the model that minimized the root mean square error (RMSE), where error was the difference between each model predicted and actual PWM command at each measured state. The resulting exponential model is shown in Equation 2. It yielded a RMSE of 0.01% PWM, meaning that the mean error of the model-predicted -PWM command to reach the desired state is 0.01%. All model identification was performed in MATLAB R2021a.

After implementing the feedfoward model, we tuned a PID feedback controller during P-Pong play to compensate for model error. The tuned control law yielded a position RMSE of 0.001 inches and velocity RMSE of 0.004 in/s over two and a half minutes of Pong play.

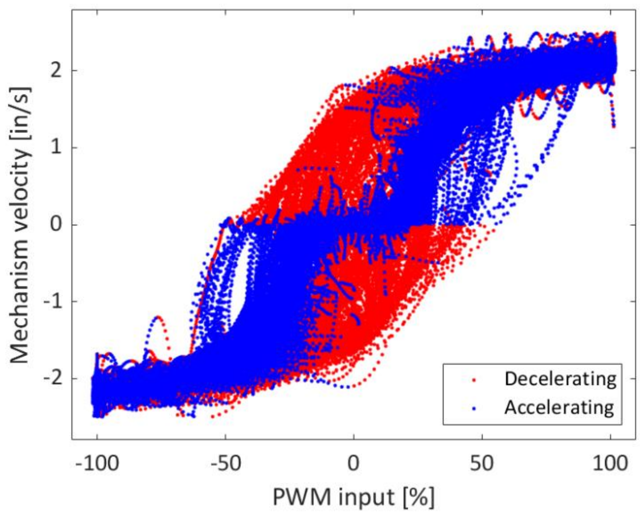

Figure 7. System identification experiment result for model-based control development.
When excited by the pseudorandom non-binary signal ranging from 0 to 100% PWM in both flexion and extension directions, the velocity output exhibited deadband and hysteresis. The hysteresis is depicted with marker colors, where increasing velocity in either direction followed a different path than decreasing velocity. Positive velocities and PWM inputs represent movements in the finger extension direction, while negative values correspond to movements in the flexion direction. The general shape of the relationship between the input and output resembles the letter "S" - small accelerations are closest to the midline "S", and as accelerations increase PWM, velocity pairs radiate outward from the "S" midline.

$$PWM = 1.032\ddot{y} + sign(\dot{y})[7.512e^{1.106\dot{y}} - 1.232]$$

Equation 2. Identified plant model.
The model takes in a desired state, linear velocity ($\dot{y}$) in in/s and acceleration ($\ddot{y}$) in in/s^2, and outputs the PWM control signal to drive the mechanism to the input state. The model accounts for hysteresis with the acceleration and sign terms.

C. Participants and experimental design

The study was approved by the UCI Institutional Review Board and subjects provided consent. We assigned 15 healthy subjects, 7 female, 2 left-handed, ages 20-51 to either the proprioceptive training (8 subjects) or control (7 subjects) group. The proprioceptive training group played P-Pong with their dominant hand in the manipulandum and their non-dominant hand holding the controller. The control group played a traditional, visually-driven version of P-Pong. Unlike for the proprioceptive training group, the ball and paddles were always displayed on screen and the dominant hand was not driven by the robot like the traditional video game. The dynamics of the ball and paddle were the same for both groups. We will refer to the proprioceptive training group as "Prop Pong" and the control group that played the video-only version as "Video Pong".

Each participant attended one session comprised of a Crisscross baseline assessment of finger proprioception acuity, 15 minutes of game play (Prop Pong for the proprioceptive training group or Video Pong for the control group), and a Crisscross post assessment. We confirmed participants understood Crisscross and their assigned Pong version with an introductory practice period. To account for differences in adapting to the rules and controls of each activity, participants practiced until they verbally acknowledged that they understood the activity. We modified the original Crisscross assessment described in [106]: instead of finger approach speed being held constant, at the start of each crossing trial a speed was independently sampled for each finger mechanism for a total of 12 trials. Each 12-trial assessment was repeated three times for a total of 36 finger crossings at both the baseline and post-training assessments. For each crossing, the robot moved the fingers from opposite flexion and extension positions towards one another with random movement start delays, and the participant indicated when they believed their fingers were overlapped by pushing a button on the handheld controller. We administered UES and IMI surveys after game play for both groups, the full UES Short Form per [109] on a 5-point Likert scale and select IMI statements on a 7-point Likert scale. For the IMI, we selected 10 statements from the four subscales Effort/Importance, Perceived Competence, Interest/Enjoyment, and Pressure/Tension [110]. We assigned subjects to their training group using an adaptive randomization technique based on their mean baseline Crisscross crossing error to attempt to match mean baseline proprioception acuity between the two groups.

D. Data analysis

We used crossing error, defined as the unsigned difference in position between the two finger mechanisms, to quantify finger proprioception acuity. For Crisscross, we calculated crossing error from the finger positions at the moment when the participant indicated that their fingers were overlapped. For Prop and Video Pong, we calculated crossing error at the moment when the ball reached the player's paddle.

We tested four research questions. First, does Prop Pong gameplay improve finger proprioception acuity? To answer this question, we evaluated the change in mean Crisscross crossing error (baseline to post) using Student's t-tests. We compared both experiment groups using the 2-sample t-test and performed 1-sample t-tests on each group to check for significant Crisscross change (versus no change). We also evaluated whether the Prop-Pong crossing error decreased over the course of game play.

Second, does Prop Pong performance predict proprioceptive acuity? To answer this question, we performed correlation to identify whether the mean of the final 36 Pong crossing errors predicted the mean of the post-training Crisscross crossing errors across subjects.

Third, is Prop Pong motivating and engaging? We evaluated engagement and motivation by comparing post-play UES and IMI survey scores of Prop and Video Pong using the 2-sample t-test. Following analysis, we converted IMI scores to the 5-point scale used for UES for reporting purposes.

Fourth, were there differences in success and number of repetitions between the groups? Given that both groups were time (not repetition) matched, and the difficulties of each game mode (Prop and Video) were inherently different, we compared overall success and repetition counts between the experiment groups using the 2-sample t-test. Repetition counts were quantified as the number of ball return attempts at the end of each 15-minute play session. Similarly, we calculated success as the total number of hits divided by the total number of return attempts at the end of the play session.

For all t-tests, we first tested for normality and variance homogeneity with the Anderson-Darling and 2-sample F-test, respectively. We performed all analyses in MATLAB R2021a.

RESULTS

A. Does Prop Pong play improve finger proprioception acuity?

Proprioceptive game play significantly reduced crossing error as robotically measured with the Crisscross assessment (one-sided, one-sample t-test, $p = 0.023$), but video-only game play did not (one-sided, one-sample t-test, $p = 0.101$) (Figure 8). The mean crossing error decreased by 0.5 ± 0.9 SD mm for the proprioceptive training group, and 2.2 ± 2.6 SD mm for the control group from baseline to post Crisscross, a difference that approached significance (one-sided, two-sample t-test, p = 0.058). In addition, the Prop Pong crossing error (proprioceptive training group) decreased over the course of gameplay (linear regression, p < 0.01 and slope -0.11 mm/min) (Figure 9).

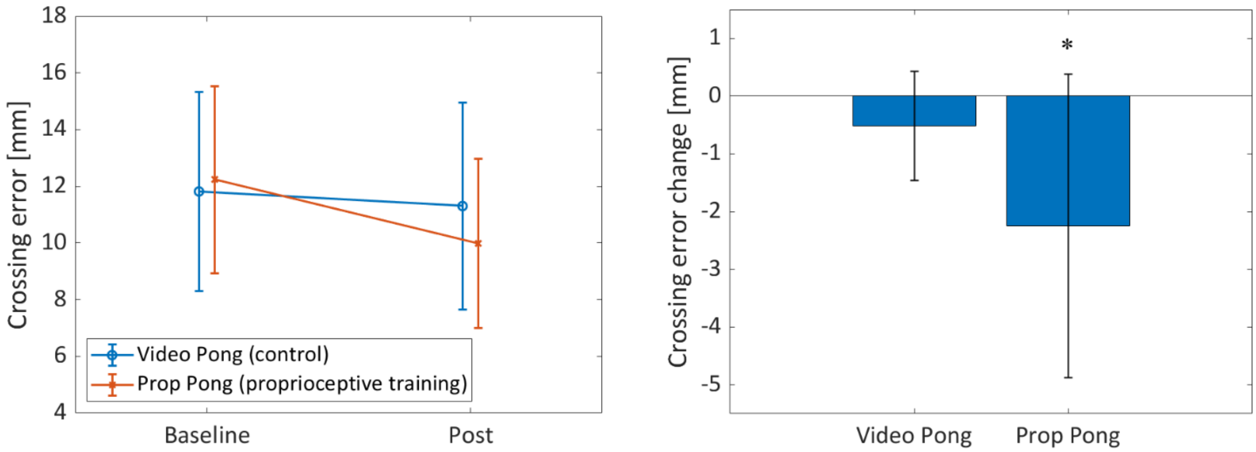

Figure 8. Finger proprioception acuity for each experimental group. Proprioception acuity was measured by the mean crossing error with the Crisscross assessment, at each timepoint (left) and change from baseline to post-gameplay (right). Error bars show ± 1 SD. * indicates significant, one-sided, t-test (p = 0.023) of change compared to zero.

B. Does Prop Pong performance predict proprioception acuity?

The Crisscross crossing errors measured post-training were not significantly correlated with the crossing-errors during the last 36 ball hits of Prop Pong play (p = 0.82).

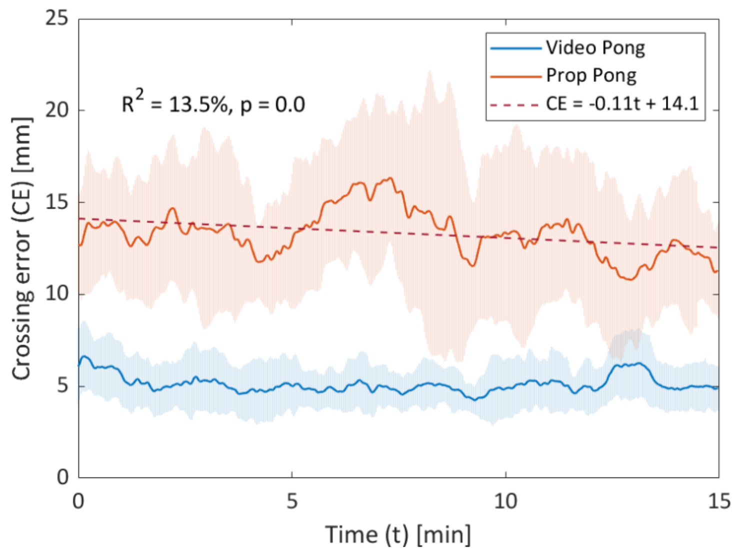

Figure 9. Mean Video Pong (control) and Prop Pong (proprioceptive training) crossing error over play time.

For each subject, we calculated a moving mean of the crossing error over a 60 sec window. Video and Prop Pong lines are means across subject moving means, shaded regions are ± 1 SD. The dashed line shows the best fit line using linear regression.

C. Is Prop Pong motivating and engaging?

Participants rated all subscales > 3.0, except IMI Perceived Competence for Prop Pong, and IMI Pressure/Tension ≤ 2.0 which is theorized to be a negative marker for intrinsic motivation [110], indicating that user engagement and experience were positive for most activities (Table 10). Scores on two of the eight subscales were significantly less for Prop Pong than for Video Pong – the Perceived Usability (UES) and Perceived Competence (IMI) subscales (2-sample t-test, $p < 0.01$).

Table 10. UES and IMI mean survey results per subscale.

Group	UES				IMI			
	FA	*PU*	*AE*	*RW*	*IE*	*PC*	*EI*	*PT*
Video Pong	4.0	4.2**	4.4	4.1	4.3	3.8**	3.7	1.7
Prop Pong	4.5	3.1**	4.2	4.4	4.6	2.2**	4.4	1.9

IMI results were converted to the 5-point scale used for UES. The UES subscales are Focused Attention (FA), Perceived Usability (PU), Aesthetic Appeal (AE), and Reward (RW). The IMI subscales are Interest/Enjoyment (IE), Perceived Competence (PC), Effort/Importance (EI), and Pressure/Tension (PT). **p < 0.01.

D. Were there differences in success and number of repetitions between the groups?

The proprioceptive training group played significantly less repetitions (two-sided, two-sample t-test $p < 0.01$) and was significantly less successful (two-sided, two-sample t-test $p < 0.01$) than the Video Pong control group. The proprioceptive training group played a mean of 192 ± 24 SD repetitions, compared to 285 ± 28 SD repetitions for the Video Pong group. The proprioceptive training group returned the ball 29.7 ± 10.1 SD % of the repetitions, while the control group returned it 74.5 ± 11.2 SD % of the repetitions (Figure 10).

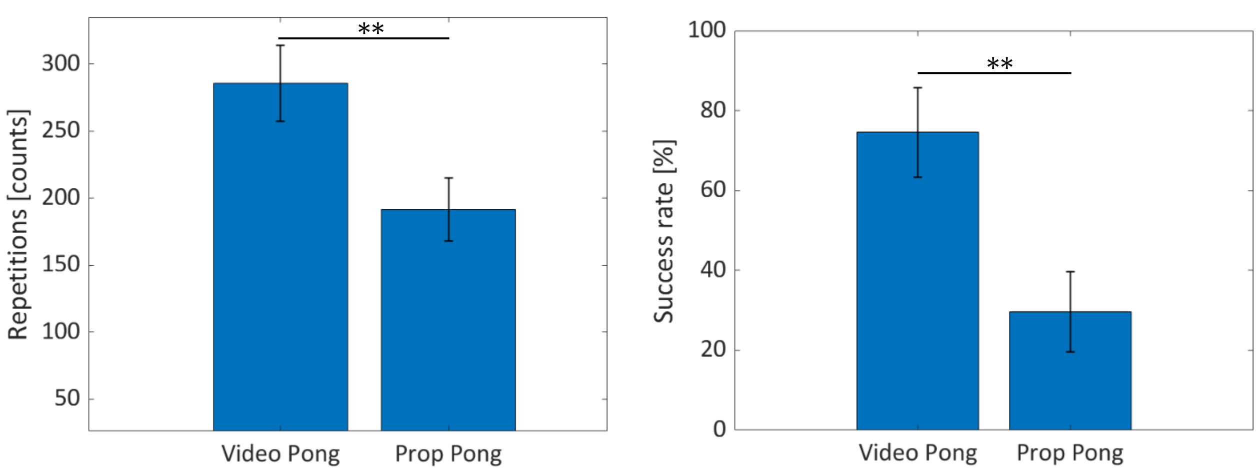

Figure 10. Differences in number of repetitions and success between groups.
Repetitions and success at the end of the 15-minute play session were compared between groups. The proprioceptive training group performed significantly less repetitions and was significantly less successful on average. Error bars show ± 1 SD. The ** indicates a significant two-sided, two sample t-test of $p < 0.01$.

Discussion

This chapter first introduced the concept of a proprioceptive computer gaming paradigm, "Propriopixels", that splits game feedback between vision and proprioception for the intense and engaging training of proprioception acuity. We then described the practical design pattern of a "binary impedance robot" for measuring and training finger proprioception. Finally, the pilot experiment with P-Pong indicated that the human sensory motor system has the ability to at least temporarily improve proprioception acuity with such game-based training.

A. PINKIE: A practical robotic device for measuring finger proprioception

Clinical techniques for measuring finger proprioception typically rely on crude tests, such as having the patient close their eyes and respond when the therapist moves their finger up or down. Currently, there are few practical robotic technologies for quickly quantifying finger proprioception. We had previously developed the FINGER exoskeleton as a way to provide high fidelity control and good backdriveablity for finger movement studies [105], [106]. We modeled PINKIE after the FINGER exoskeleton [105] in that it incorporates mechanisms for a furling/unfurling motion of the index and middle fingers. Unlike FINGER, however, we simplified the mechanism to a finger-contacting roller and linear slide instead of an 8-bar linkage, we used low-cost actuators and sensors, and we fabricated custom components with 3D printing and laser cutting. Also unlike FINGER, PINKIE has the disadvantage that it can only implement two finger mechanism impedances: either low impedance, in which it is very backdriveable but can't actively drive the finger, or high impedance, in which it is not user backdriveable but can actively drive the finger. We made the impedances switchable via a simple, automatic, magnetic clutch. While this design limits the capability of PINKIE to provide active assistance, these are the two impedance modes essential for most proprioceptive testing paradigms.

B. P-Pong: Implementing the Propriopixels proprioceptive computer gaming paradigm

While games have been developed for proprioception training [38]–[40], previous approaches do not integrate proprioception as a required input to the player's decision making process. Here we presented the concept of a proprioceptive computer game. In the Propriopixels paradigm we propose, players must continuously make game decisions based on sensed finger position as the game robotically moves the fingers, instead of screen pixels, to express the motion of game elements. We implemented this paradigm on the classic arcade game Pong, but it could be applied to many other existing video games. Our intent is to use Propriopixels to increase the intensity and motivation for proprioceptive training.

C. Pilot Results with Proprioceptive-Pong

Playing 15 minutes of Prop Pong caused a significant reduction in Crisscross crossing error, a measure of finger proprioceptive acuity, while playing Video Pong did not. This indicates that finger proprioceptive acuity can be improved at least temporarily using a proprioceptive gaming paradigm. We also found that Prop Pong crossing error decreased over time, which may be due in part to improved proprioception, although other factors influence crossing error during Prop Pong gameplay as well (see below).

We have shown previously that crossing error measured with the Crisscross test is sensitive to the proprioceptive decline known to occur with aging [106] and is predicted by a combination of neural function (connectivity between ipsilesional secondary somatosensory cortex and ipsilesional primary motor cortex) and neural injury (total sensory system injury) [111]. In addition crossing error measured at baseline predicts the amount of functional benefit attainable with robotic finger training [41], [71]. The finding that crossing error can be reduced with training is intriguing, given the significance of this marker of proprioception.

Surprisingly, we did not find a significant correlation between Crisscross and Prop Pong crossing errors across subjects. This may be because our sample size was small. It may also be because additional physiological mechanisms influence Prop Pong play performance. Unlike Crisscross, Prop Pong required interhemispheric coordination, was dependent on non-dominant hand motor control ability, and involved controlling an unfamiliar entity (i.e. the "paddle" finger, for which the transfer function between the joystick input and paddle finger movement exhibits nonlinear dynamics). Also, unlike in Crisscross, in Prop Pong players could potentially use visual information to recalibrate their proprioceptive estimates of their finger positions [112]. Specifically, each time the ball and player paddle appeared on screen (which was only while the ball was traveling away from the player paddle), the player could use simultaneous visual and proprioceptive information to update their belief of their current finger position. One implication of the non-significant correlation between Crisscross and Prop Pong crossing errors is that it may be better to define and implement a dedicated finger proprioception assessment (such as the Crisscross Test), rather than trying to infer proprioception acuity from gameplay metrics, as the latter involves other physiological mechanisms that add noise to proprioception assessment.

Participants rated Prop Pong positively for all queried engagement and intrinsic motivation subscales, except the Perceived Competence subscale of the IMI for the Prop Pong group (see below). While this result may seem unsurprising since Prop Pong is based on a successful and ubiquitous game, we changed the input-output structure through which the player interacted with game elements, which is a key ingredient of the player's experience. We also found small but significant differences between Prop and Video Pong subscores on some rating elements that most likely are attributable to the input-output structure. Participants scored Prop Pong significantly lower than Video Pong on both the Perceived Usability UES subscale and Perceived Competence IMI subscale. The UES subscale queries whether the activity was frustrating, confusing, and taxing, and the IMI subscale probes feelings related to performance satisfaction and skill level. Indeed, the Prop Pong group was not only significantly less successful than the video-only group, their success rates were low with a mean of 29.7% - they missed 7/10 times on average. As the input-output structure was the defining difference between Prop and Video Pong play, it could be that the combination of moving the paddle with the non-dominant controller hand, sensing movement with the manipulandum hand, and tracking remaining game elements on screen introduced cognitive burden that was somewhat more confusing and taxing, which in turn led to decreased performance. Perhaps with a lower, tunable game difficulty players would rate usability and competence higher.

The Prop Pong group had significantly lower success and repetitions than the Video Pong group. Both factors have been shown to affect learning [45]–[47]. Through a computational modeling approach, Wilson et al. established 75%-85% success as an optimal rate for learning [113]. And while the Video group was successful approx. 75% of the time on average, the Prop Pong group was far below at 30%. Furthermore, since Prop Pong players restarted the game many more times due to their low success, they only attempted to return the ball 192 times on average, whereas the Video Pong group attempted 285. Together these factors may have reduced the potential training benefit of Proprioceptive-Pong play.

Key directions for future work are to regulate Proprioceptive-Pong success rate to an optimal 75%-85% range for learning, and in an experimental setting, match both repetitions and success between groups. The effects of different control scheme designs on game difficulty and proprioceptive improvement should also be considered. A simplified alternative scheme could be to replace the joystick held in the contralateral hand with an unactuated, stripped-down version of the manipulandum, and a "mirror-match" control law where the manipulandum matches the position of the paddle finger with the position of the contralateral input finger.

Conclusion

These results support the potential value of the proposed proprioceptive computer gaming paradigm Propriopixels for improving finger position sense, and of the simple binary impedance robot PINKIE for delivering proprioceptive assessments and gamified training. Our goal is to translate that value to different rehabilitation settings, including the clinical setting where finger proprioception assessment is crude and few methods are available for intense proprioceptive training, and the home setting where rehabilitation adherence is low.

Portions of this work were previously published in the 2021 International Conference of the IEEE Engineering in Medicine & Biology Society [114].

CHAPTER 4: THE EFFECTIVENESS OF GAME-BASED ROBOTIC AND NON-ROBOTIC PROPRIOCEPTION TRAINING

Summary of the Chapter

Proprioception deficits are common after stroke. Proprioception is responsive to training and has been identified as a key input to motor learning. Yet, few engaging, intense methods exist for proprioception training and to our knowledge none exist for the fingers. We investigated the potential benefits of two types of gamified proprioception-targeting training with an updated version of the Proprioceptive-Pong game, which we enhanced with several refinements: targets and leveling intended to increase engagement with the game, a new non-robotic proprioception-targeting training mode, and an automatic difficulty control algorithm to titrate success rate. Then, in an experiment with 36 unimpaired participants, we compared the effects of three types of training, (1) Propriopixels and (2) Visioception, which targeted proprioception by requiring the use of proprioception to make in-game decisions (the first type robotic and the second non-robotic) and (3) a typical video-only version of the game. Our outcome measures were changes in in-game performance, Crisscross crossing error (a measure of passive proprioception acuity), the standard Box and Blocks Test, and a blindfolded version of the Box and Blocks Test across three time points (baseline, post, and short term follow up) using repeated measures analysis of variance (RM ANOVA). Across all outcome measures, there were no significant interaction effects between group and time point. P-Pong game error, standard BBT, and blindfolded BBT significantly improved over time ($p < 0.001$). Crisscross crossing error however did not significantly change over time ($p = 0.36$), with small <1 deg fluctuations in crossing error across groups and time points. We investigated potential causes of this result and found that Crisscross crossing errors were significantly lower ($p < 0.01$) than the previous chapter at baseline

(where we measured significant Crisscross improvements), a finding that we discuss. Next, we evaluated the performance of the success control algorithm and found that it drove participants to the desired 80% success rate with a final rate of 79.1 ± 3.2 SD % across all three training groups completely virtually without physically assisting participants. We verified its time series convergence using nonlinear regression and found that an exponential decay model explained 42% of the variance in success rate errors over the duration of P-Pong play. Lastly, we investigated the effects of varying approach speed during the Crisscross assessment using linear regression and found that crossing error was not significantly correlated with approach speed. Lastly, we discuss these findings and future directions for the investigation of gamified proprioception-targeting training.

INTRODUCTION

In the last chapter we presented a novel proprioceptive computer gaming paradigm that we call "Propriopixels" and a simple robot design pattern that we call a "binary impedance robot" for playing Propriopixels games. We evaluated these developments in a pilot study of unimpaired participants and found that 15 minutes of play improved passive proprioception acuity, which we measured robotically. In this chapter, we further investigated the potential benefits of gamified proprioception targeting training with an enhanced version of the Proprioceptive-Pong (P-Pong) game.

Proprioception and movement are closely linked. The motor system uses proprioceptive information to plan [115], [116] and control movement [117]–[119], and proprioceptive training has been shown to improve motor learning [120]. In a recent review, Winter et al. found that across 50 studies that investigated active movement and balance interventions, improvements in proprioceptive outcomes were comparable to improvements in motor outcomes. Interestingly, studies that instead trained somatosensation without movement during training tended to report higher gains in proprioceptive than motor outcomes [121]. Further, they found that somatosensation targeting training, such as with somatosensory discrimination or stimulation had small to medium effect sizes, suggesting that propriomotor training may be more effective for improving proprioceptive and motor performance than somatosensory training alone.

In this chapter we built on our previous work by attempting to increase the training benefit of our Propriopixels paradigm. We enhanced the P-Pong game by adding targets and leveling with the intension of increasing player engagement, a new non-robotic proprioception targeting game mode that we call "Visioception", and an automatic difficulty control algorithm to titrate success rate. We evaluated these refinements by comparing three types of active finger movement training. Each type varied the ways in which game elements were conveyed: constantly via a screen ("Video-only", like a video game), multimodally via a screen and/or a robot-driven finger mechanism (Propriopixels), or intermittently via a screen (Visioception). Specifically, we evaluated the effect of these three versions of target-based P-Pong play on (a) in-training performance, (b) proprioceptive acuity measured by Crisscross, and (c) hand dexterity measured by two versions of the Box and Blocks Test (BBT) – a standard and blindfolded version.

We hypothesized that all groups would decrease P-Pong finger position errors that we refer to as "crossing errors" during training. We hypothesized that targets would increase crossing errors for the Propriopixels and Visioception groups since for both groups the paddle was not displayed on screen (while the ball approached) and the desired paddle contact area locations were novel. For example, for the Propriopixels group a top target cued players to move their index paddle finger a distance above their middle ball finger, however the distance was unknown at the start of play. Following results from the previous chapter, we hypothesized that P-Pong training would significantly improve Crisscross crossing errors for all groups due to the proprioception training effects of active movement noted in previously studies– all groups controlled the P-Pong paddle with active index finger movements unlike the previous chapter. For the effect of training on hand dexterity, we hypothesized that no group would improve on the standard BBT assessment because it relies on vision and that the Propriopixels group would improve most on the blindfolded assessment because Propriopixels training would have the greatest effect on active position sense and that the blindfolded assessment performance would depend more strongly on active position sense than the standard assessment.

Lastly, we aimed to (d) verify that the success rate control algorithm is capable of regulating success across all groups and (e) investigate the relationship between Crisscross crossing error and finger movement speed. We hypothesized that the success control algorithm would regulate player success across all groups and that Crisscross crossing error would increase with speed. We next present updates to the P-Pong game, our experimental protocol, and our analysis pipeline.

METHODS

A. *Target-based Proprioceptive-Pong and its game modes*

Following the study in the previous chapter, we made several updates to the Proprioceptive-Pong (P-Pong) game design which we describe here.

We fixed the horizontal ball speed (it previously ramped from a low initial speed to a high-speed during game play) and the number of returns (20) per match to fix the duration and number of repetitions per match for experimental control purposes. With this change the only consequence of missing was a lower score.

We added "pseudo" leveling and high scores to increase attention and further motivate players. There was no actual change in difficulty from level to level – difficulty was adjusted algorithmically per return by the success control algorithm based on the player's performance (more difficult for each hit, less difficult for each miss). Players were automatically leveled up whenever they achieved the desired success rate set for the success control algorithm, meaning all players were leveled up at similar and controlled rates. From the player's perspective, they leveled up whenever they achieved a score corresponding to the desired success rate, e.g. for a success rate of 80% they needed 160 points at 10 points per hit to level up, which again, was regulated by the success control algorithm. We developed a pseudo high scores list based on feedback from participants in Chapter 3 – several asked how well they were playing and or commented about it in some way. After each match, the game generated a high scores list with fake player names and scores. The fake names and scores regenerated each time the player started a new level.

In an exploratory venture intended to test if we need a robot for propriomotor training, we added a non-robotic proprioception-targeting training mode that we call "Visioception". A hybrid of the Video-only and Propriopixels modes, it requires players to combine visual and proprioceptive information to estimate where the paddle should be located to get a hit. Like the Video-only mode it uses the robot as a sensor to measure paddle finger movement while the middle finger is stationary (the ball state is always displayed on screen), and like the Propriopixels mode the current paddle position is not shown on screen while the ball approaches, requiring the player to estimate without vision where they are moving their paddle.

To further stress proprioception we reformulated P-Pong's measure of success. Instead of conditioning success on whether the player returns the ball (as in the previous game iteration and the classic video game), with targets the player's goal is to hit the ball to one of three locations on the opposing wall. A player hits a target by contacting the ball with one of three respective zones on their paddle. In Propriopixels mode this corresponds to how the index paddle finger must be positioned relative to the middle ball finger. Previously Propriopixels could effectively be played without visual information – the player could constantly match the position of their middle finger with their index finger as the centers of the ball and paddle vertical locations corresponded to the positions of their middle and index fingers, respectively. Instead, the visually displayed targets encode whether the paddle finger should be above (top target), matched (middle target), or below (bottom target) the ball finger, requiring the player to decide where to position their index finger based on both proprioceptive ball information and visual target cues. In other words, if the target were held constant at the middle target then the game would be relatively unchanged, and as the target varies the player must vary the orientation of their index and middle fingers accordingly. A new target (top, middle, or bottom) was randomly sampled from a uniform distribution each time the ball began moving toward the player's paddle. We depict and further describe both the targets and all game modes in Figure 11. Screenshots of each game mode are in Figure 12. Lastly, we implemented an algorithm to regulate player success rates, which we describe in the following sub section. Regulating game success is an important goal as previous work suggests there is an optimal game success level that encourages motivation and self-efficacy during robotic and sensor-based training [71], [122].

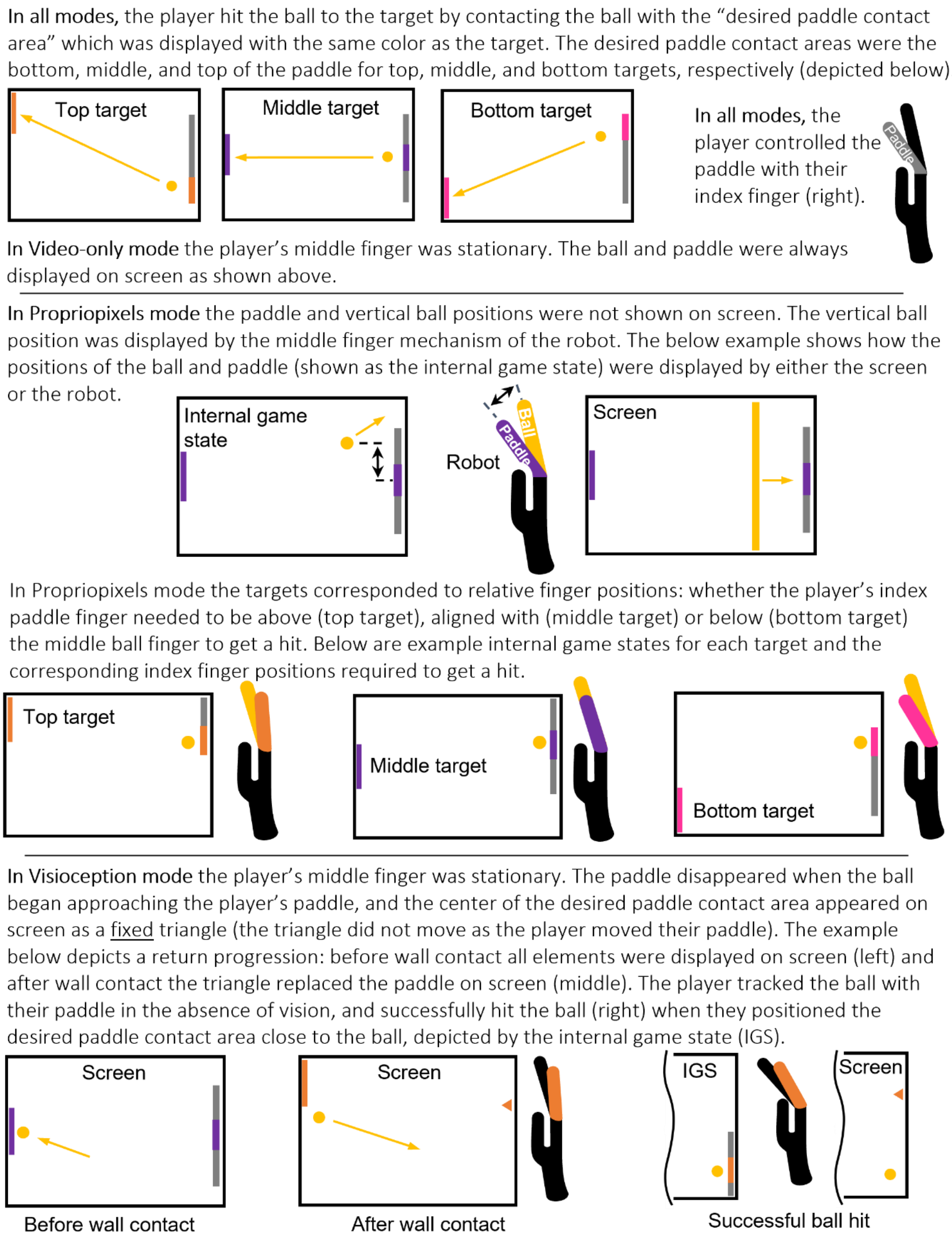

Figure 11. Explanation of Proprioceptive-Pong targets and display modes.

Instead of just returning the ball, the player's goal was to hit the ball to the target displayed on the opposite wall (top). The three game modes, Video-only (top), Propriopixels (middle) and Visioception (bottom) changed how vertical ball and paddle positions were conveyed.

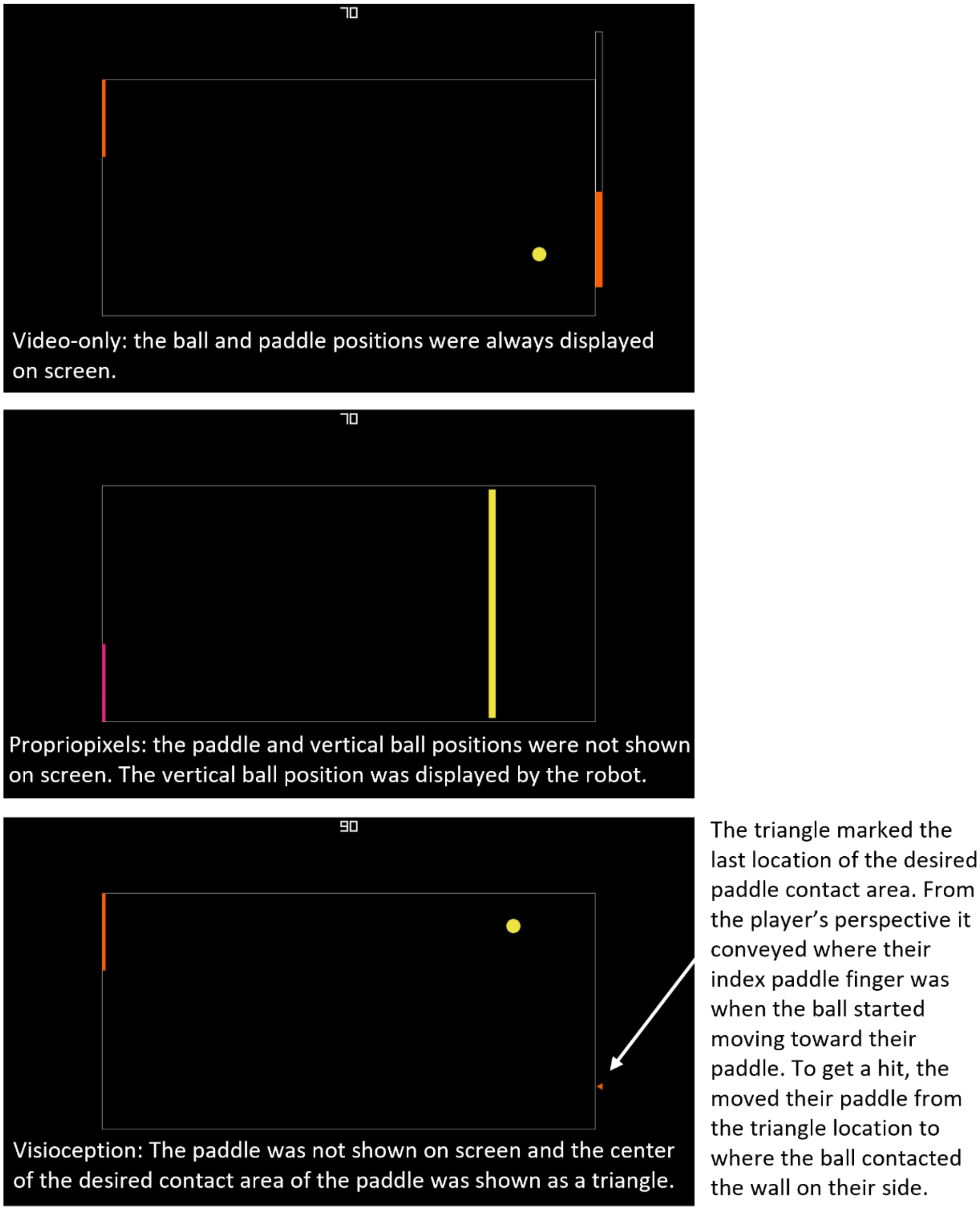

Figure 12. Screenshots of each P-Pong mode as the ball moves toward the player's paddle. In each mode, the current target was displayed on the left and the player's paddle was on the right. When the player hit the ball with the desired paddle contact zone then the ball moved to the target. We hid the AI paddle to reduce visual clutter around the targets since they acted as cues for making gameplay decisions. For the Propriopixels (middle) and Visioception (bottom) modes, the ball and/or paddle were not all displayed on screen while the ball moved toward the player's paddle. When the ball moved in the opposite direction away from the player's paddle, the ball and paddle were both shown on screen like the Video-only mode (top).

B. Real-time virtual success control

In the previous chapter we found that there was a significant difference in success rates between the Video-only and Propriopixels groups: 74.5 ± 11.2 SD % vs 29.7 ± 10.1 SD %, respectively. This may have reduced training benefit in the Propriopixels group as they fell below the optimal 75%-85% range for learning proposed by Wilson et al. [113]. Baranes et al. also that participants started with easy games then progressed to more difficult ones and repeated moderate to high difficulty games frequently in a game-based self-exploration task [123]. The difficulty of the Propriopixels game therefore may have begun and remained too challenging. We implemented an updated version of the success control algorithm [124], which has been used to regulate success in the Guitar Hero-like robotic rehabilitation game "Rehab Hero" [71], [105].

We made a simple modification to the original algorithm to enhance its ability to down regulate success. In practice, the original algorithm adjusted robot Proportional-Derivative control gains to assist players along a "successful" movement trajectory. For stroke survivors the initial game difficulty was often too hard – they were not capable of completing the movements required to be successful at high success rates. Therefore, as they played the algorithm incremented control gains to assist players along a trajectory until they were able to perform movements at the preset desired success rate. In this formulation however, once the control gains are minimized the system is no longer controllable – the control signal saturates and can no longer decrease success rates (unless we were to program the robot to resist players). This saturation behavior was observed in unimpaired participants in [105]. We added a second region to the algorithm that we call the "challenge" region; it has the same structure of the original algorithm that we now call the "assist" region. In the assist region parameters are varied that are capable of upregulating success to the desired rate, while in the challenge region parameters are varied that are capable of downregulating success for a range of player abilities (Equation 3).

In the study described in this chapter, we varied only virtual game parameters to control success, meaning that we did not use any physical robotic assistance to control success rates. In the "assist" region, the algorithm varied the paddle height, where decreases in the assistance gain (k_a in Equation 3) were mapped to proportional increases in paddle height. In the "challenge" region, the algorithm varied both paddle height and vertical ball speed – increases in the challenge gain (k_c in Equation 3) were mapped to proportional decreases in paddle height and increases in vertical ball speed. Although in this implementation varying only the paddle height is theoretically adequate – if the paddle height were minimized then success rates could be down regulated to 0%, this two-region algorithm version was also useful for balancing the game design. "Balance" is an important concept in game design theory and practice [125]–[127]. Quoting a recent review by Becker and Görlich, "In a highly abstract way one can interpret 'game balancing' as the activity of tuning a game's rules, levels, difficulty, numbers, algorithms etc. to achieve desired goals…such as keeping a game winnable, making it fair for all players, keeping it challenging, making it replayable, etc." [128]. We tended to lose attention when the ball speeds were too low and found that down regulating success with miniscule paddle heights was frustrating and seemed like an unwinnable task. Using vertical ball speed together with paddle height to down regulate success in the challenge region required smaller decreases in paddle height that brought a motivating challenge to the game.

Success control algorithm $(p_\infty, \delta_a, \delta_c)$

1. initialize difficulty $k_d = k_a + k_c = 1 + 0$ where $k_a, k_c \in [0,1]$
2. for each attempt i do
3. present difficulty to player and obtain success result
$$s = \begin{cases} 0 & successful \\ 1 & unsuccesful \end{cases}$$
4. compute challenge indicator to select challenge or assist region
$$I_c = \begin{cases} 0 & k_d^i < 1 \; or \; \left(k_d^i = 1 \; and \; s = 0\right) \\ 1 & else \end{cases}$$
5. update difficulty $k_d^{i+1} = k_c^i + I_c(s\delta_c - (1-s)\alpha\delta_c) + \; k_d^i + (1- \; I_c)(s\delta_d - (1-s)\alpha\delta_c)$
where $\alpha = \dfrac{p_\infty}{1 - p_\infty}$
6. *end for*

Equation 3. The success control algorithm.

This algorithm operates under the same statistical principle as the original, that the actual success rate p_a converges to the desired success rate p_∞ over infinite time. On every success or failure the difficulty (parameterized by the gain k_d) is stepped: on a success it is incremented by a fixed step size δ and on a failure it is decremented by a fixed ratio of the step size $\alpha\delta$, where α is a function of p_∞ and δ is tuned a priori. For example, an 80% desired success rate $p_\infty = 0.8 \Longrightarrow \alpha = 4$. When $p_a = 0.8$, there will be four successes for every one failure, meaning that over five attempts the net change in difficulty will be zero: $k_d^i = \; k_d^{i-5} + 4\delta - 4\delta = \; k_d^{i-5}$, where $+4\delta$ corresponds to the four successes and -4δ to the one failure. Meaning, that although the difficulty is constantly being modulated (it is not stationary when $p_a = p_\infty$), it is modulated around a difficulty that yields the desired success rate. Of note, the algorithm has no explicit knowledge of p_a (it is not in the algorithm) as the history of successes and failures is not tracked – the only "feedback" measurement is the success outcome at the current time. Historical success information is stored in the difficulty gains.

Different from the original algorithm, there are two regions "assist" and "challenge". It only operates in one region at a time, meaning that the challenge gain k_c is not modulated until the assistance gain k_a is saturated. This allows different parameters of the activity or game to be modulated depending on how well the participant is performing. To ensure convergence, we recommend the assistance gain be theoretically capable of upregulating success to 100% when minimized (lower gains correspond to less difficulty), and the challenge gain to be capable of downregulating success to 0% when maximized. Note that setting $I_c = 0$ recovers the original algorithm, as when $I_c = 0, k_c = 0$ per the initial condition and I_c definition in step 4.

C. Participants and experimental design

This study was approved by the University of California, Irvine Institutional Review Board and participants provided consent. We assigned 36 participants (18 female, 6 left-handed, ages 18 to 35 with 24.4 ± 4.0 SD years) to one of three training groups. Each training group played a different mode of P-Pong, either Video-only, Propriopixels, or Visioception. We evaluated sensorimotor capacity with three assessments: the standard Box and Blocks Test (BBT), a blindfolded version of the standard BBT, and the Crisscross assessment (the robotic assessment of passive finger proprioception presented in the previous chapter), each measured at three time points: baseline (before training), post training, and at a short term follow up session (Figure 13).

We made three changes to the robotic activities (P-Pong and Crisscross) in addition to those detailed previously in this section. First, all robotic activities were carried out using the FINGER robot [105], as opposed to PINKIE. There are several key differences between the two robots, but at a high-level FINGER is considerably more complex. FINGER can render a continuum of mechanism impedances using high bandwidth actuation and sensing which we have used to implement high fidelity assist-as-needed control algorithms. In this study, however, we used FINGER in the same way as PINKIE where it did not physically assist players – all assistance was performed with virtual game parameters as described in the previous sub section. Second, we modified the FINGER Crisscross assessment to test a range of finger velocities. At the start of each trial when each finger was separated (one flexed and one extended), the fingers paused for a randomly sampled time duration then approached one another at the same speed. An additional trial was then performed at the same speed. The speeds were randomly sampled without replacement from a set of 10 unique speeds, yielding a total of 20 trials per assessment. Third, we changed the paddle control scheme to require active index finger movement. In the previous chapter participants controlled their index paddle finger via a joystick held in their contralateral hand (outside the robot), i.e. the index finger was passively driven by the robot. We modified all game modes to use an active index finger control scheme.

In all three game modes player paddle was controlled with the index finger, and the game mode varied how the vertical ball and paddle positions were conveyed, as described above. To summarize briefly, in the Video-only mode the middle finger was stationary, and all game elements were displayed on screen like a typical video game. In the Propriopixels mode, the vertical ball position was displayed by the robot on the player's middle finger and the visual information was varied per the description in Figure 11. In the Visioception mode the middle finger was stationary, and the ball was always displayed on screen like the Video-only mode. Unlike Video-only, the paddle position was not updated on screen whenever the ball moved toward the player's paddle as in Figure 11. All modes were played with targets. We did not tell participants about the success control algorithm, that the levels had no effect on difficulty, nor that all the names and scores in the high scores list except their own were generated algorithmically.

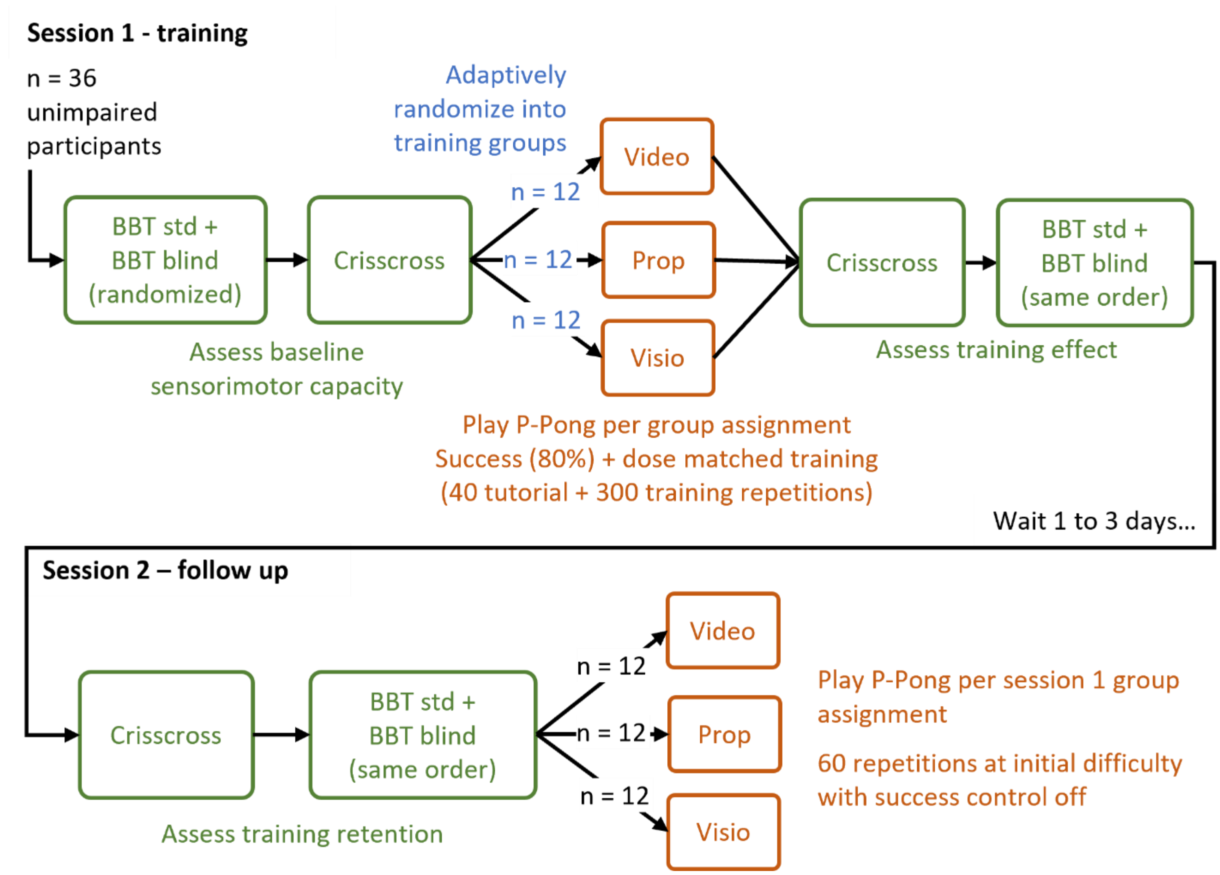

Figure 13. Study data collection procedure.

The study consisted of two sessions comprised of three assessments and one gamified sensorimotor training activity. During session 1, the assessments, the standard Box and Blocks Test (BBT), a blindfolded version of the BBT, and the Crisscross assessment were administered before and after training. Participants were randomized into one of two BBT test groups to compensate for test ordering effects: performing the standard or blindfolded test first. Following baseline Crisscross, participants were adaptively randomized into one of three Proprioceptive-Pong training groups, Video-only (Video), Propriopixels (Prop), or Visioception (Visio), using the mean and standard deviation of their crossing error. Participants first played a tutorial comprised of two matches, one without targets and one with targets, then played 15 training matches (300 repetitions). The ball horizontal speed was held constant to match training time across participants, and every participant began with the same initial game difficulty. During the second tutorial match, the success control algorithm was turned on, and the difficulty automatically modulated for the remainder of the training session with an 80% desired success rate. The assessments were repeated following training. One to three days later, participants attended a follow up session. They repeated the assessments, then played three P-Pong matches (60 repetitions) at the initial difficulty with the success control algorithm off. The BBT versions were always performed in the same order that was randomized at the start of the data collection.

D. Outcomes

For P-Pong and Crisscross, we used the mean of the unsigned difference between the positions of the index and middle fingers (Crisscross) or difference between the center of the desired paddle contact area and vertical position of the ball at paddle contact (P-Pong), which we refer to as "crossing error" as our primary outcomes. Crossing error is expressed as a metacarpophalangeal joint position error in degrees. For Crisscross crossing errors, we observed outliers during the first trial of several assessments of multiple participants. We therefore did not analyze the first trial of any assessment, leaving a total of 19 analyzed trials per assessment.

For both the BBT standard and blindfolded assessments, we used the number of blocks moved in one minute as our primary outcome.

For the success control algorithm, we used success rate error as our primary outcome. We computed success rates for each participant as the total number of hits divided by the total number of attempts (the sum of hits and misses), starting from when the success control algorithm was turned on during the tutorial.

To analyze the effect of finger speed on Crisscross crossing error, we used crossing errors from individual trials and the unsigned difference of the speeds, which we refer to as "approach speed" $s_a = abs(v_i - v_m)$, where v_i, v_m are the velocities of the index and middle fingers, respectively. Note that as the fingers are always moving in opposite directions (toward or away from one another) by design of the Crisscross assessment, the velocities are always opposite it sign, and "approach speed" is indeed the speed at which the fingers approach each other.

E. Statistical analysis

For all activities (Crisscross, P-Pong, and BBT standard, and BBT blindfolded) we tested for training and retention effects in the primary outcome using repeated measures (RM) analysis of variance (ANOVA) over three time points. For Crisscross and BBT, we used assessments from baseline pre-training, post training, and follow up time points. For P-Pong we used crossing errors from the first match following the tutorial as our baseline time point, the last match of the training session as our "post" time point, and the last match of the second session as our "follow up" time point. We tested all the outcome measures for normality using the Anderson-Darling test. The crossing errors were non-normal and we transformed them using the within-subjects z-score transformation described in [129]. We performed the RM ANOVA using the built in "ranova" function from MATLAB 2021a, which automatically checked for sphericity and corrected for sphericity violations. We included two terms in each RM ANOVA model, a time point within-subjects term and experimental group between-subjects term. We corrected for multiple comparisons using Fisher's least significant difference procedure, where for each term with a significant p value < 0.05, we compared the groups within that term using two-sample t-tests.

To evaluate the effects of targets on P-Pong crossing error, for each group we compared top to middle target crossing errors and bottom to middle target crossing errors at each time point using two-sample t-tests. As the crossing errors were non-normal, we compared within-subjects z-score transformed crossing errors. For example, using crossing error z-scores from the first training match, we compared top to middle target crossing errors and bottom to middle target crossing errors for the Video-only group.

To evaluate the performance of the success control algorithm, we used nonlinear regression to fit an exponential function to the success rate errors and verify that the actual success rates were converging to the desired 80% success rate.

To investigate the effect of approach speed on Crisscross crossing error, we used linear regression to fit a crossing error versus approach speed model for each participant using the baseline assessment only. We corrected for multiple comparisons using the false discovery method presented by Yekutieli and Benjamini [130]. We performed all analysis in MATLAB 2021a.

RESULTS

A. *Training effects on P-Pong crossing errors*

Crossing errors were initially lowest for the Video-only group (2.3 ± 2.0 SD deg), and similar for the Propriopixels (7.8 ± 6.4 SD deg) and Visioception group (6.2 ± 6.2 SD deg). The crossing errors decreased over time for all groups ($p < 0.01$), confirming our hypothesis that learning would occur. During the last match of the follow up session, the crossing errors were 1.4 ± 1.1 SD deg, 5.0 ± 3.9 SD deg, 4.6 ± 4.6 SD deg, for Video-only, Propriopixels, and Visioception, respectively (Table 11).

We evaluated the effects of time and interactions between time and experimental group using RM ANOVA. The RM ANOVA time term was significant ($p < 0.01$), while the experimental group-time interaction term was not significant ($p = 0.17$), indicating that all groups significantly decreased their crossing error during game play (Table 11). There were significant differences between all the time points ($p < 0.001$ for all three time point combinations), indicating that the observed decreases in mean crossing errors from baseline to post to follow up were significant.

Table 11. Summary of outcomes for P-Pong, Crisscross, and BBT activities and hypothesis testing results.

	Grp	Mean ± 1 SD deg of crossing error (P-Pong and Cc) or BBT score: Base	Post	Follow	F, p RM ANOVA: Time	Grp: Time	p between Time points: Base to Post	Post to Follow	Base to Follow
P-Pong	Vid	2.3 ± 2.0	1.9 ± 1.5	1.4 ± 1.1	44.3, 0.00**	1.8, 0.17	0.00**	0.00**	0.00**
	Prop	7.8 ± 6.4	5.3 ± 5.3	5.0 ± 3.9					
	Vis	6.2 ± 6.2	4.6 ± 4.5	4.6 ± 4.6					
Cc	Vid	4.9 ± 1.5	5.6 ± 1.7	4.9 ± 1.7	0.8, 0.36	1.1, 0.34	NC	NC	NC
	Prop	4.6 ± 0.8	4.8 ± 1.0	5.2 ± 1.5					
	Vis	4.6 ± 1.8	4.6 ± 1.2	4.6 ± 2.0					
BBT Std	Vid	69.1 ± 7.0	76.1 ± 8.3	72.8 ± 7.9	14.8, 0.00**	1.6, 0.19	0.00**	0.00**	0.00**
	Prop	75.7 ± 5.4	78.7 ± 5.7	78.6 ± 6.4					
	Vis	71.1 ± 7.2	77.0 ± 9.4	71.8 ± 6.6					
BBT blnd	Vid	49.3 ± 7.6	54.1 ± 7.1	51.8 ± 7.9	28.6, 0.00**	0.4, 0.65	0.00**	0.00**	0.00**
	Prop	52.3 ± 5.5	56.5 ± 5.1	55.0 ± 5.2					
	Vis	50.4 ± 3.4	56.7 ± 5.8	54.1 ± 4.7					

**p < 0.01. F-statistics and p-values are listed for the RM ANOVA per model term. Time point comparison two-sample t-test p values are listed for each time point combination. Abbreviations: P-Pong, Proprioceptive-Pong; Cc, Crisscross; BBT Std, standard Box and Blocks Test; BBT blnd, blindfolded Box and Blocks Test. SD, standard deviation; F, F-statistic; p, p-value; RM ANOVA, repeated measures analysis of variance; Grp, experimental group; Vid, Video-only group; Prop, Propriopixels group; Vis, Visioception experimental group. Time, time point; Grp:Time, group-time point RM ANOVA interaction term; Base, baseline time point (first training match for P-Pong); Post, post time point (last training match for P-Pong); Follow, follow up time point (last match of follow up session for P-Pong); NC, not calculated due to non-significant time p value in RM ANOVA.

Concerning target dependent errors, the Video-only group had small, non-significant variations at the start of training across the top, middle, and bottom targets with crossing errors of 2.3 ± 2.0, 2.2 ± 2.1, and 2.5 ± 2.0 SD deg, respectively ($p = 0.85$ and $p = 0.35$ for top to middle and bottom to middle comparisons, respectively). In contrast, the Propriopixels group middle target crossing errors were significantly different from the top target crossing errors ($p < 0.01$) and approached a significant difference from the bottom target crossing errors ($p = 0.05$), with crossing error means of 9.2 ± 7.4, 5.8 ± 3.5, and 8.4± 7.2 SD deg for the top, middle, and bottom targets, respectively. The Visioception group middle target crossing errors were significantly different from the bottom target crossing errors ($p < 0.01$), but not from the top target crossing errors ($p = 0.12$) with mean crossing errors of 6.0 ± 6.6 SD, 4.7 ± 4.7 SD, and 7.8 ± 6.9 SD deg, for the top, middle, and bottom targets respectively. This confirmed our hypothesis that targets would only increase error for the Propriopixels and Visioception groups. Interestingly, this variation across targets decreased from start of training to follow up for the Propriopixels group: there was no significant difference in crossing errors between the middle and top nor middle and bottom targets at follow up ($p = 0.60$, $p = 0.66$, respectively). For the Visioception group, where there was no significant difference between the top and middle and bottom and middle targets at training end ($p = 0.23$, $p = 0.06$, respectively), however at follow up there was a significant difference between the top and middle target crossing errors ($p < 0.01$). The p values for all target comparisons are listed in Table 12. The mean trajectory of crossing errors per target are plotted in Figure 14 and crossing errors per time point are plotted in Figure 15.

Table 12. Summary of P-Pong crossing errors per target and hypothesis testing results.

		Mean ± 1 SD deg of crossing error (CE) and p value comparisons					
		Training start		Training end		Follow up	
Group	Target	CE [deg]	p	CE [deg]	p	CE [deg]	p
Video-only	Top	2.3 ± 2.0	0.85	1.9 ± 1.6	0.72	1.2 ± 0.9	0.20
	Middle	2.2 ± 2.1		1.9 ± 1.3		1.5 ± 1.3	
	Bottom	2.5 ± 2.0	0.35	1.8 ± 1.6	0.32	1.5 ± 1.1	0.75
Propriopixels	Top	9.2 ± 7.4	**0.00****	5.5 ± 5.1	0.18	5.1 ± 3.9	0.60
	Middle	5.8 ± 3.5		4.5 ± 3.0		5.0 ± 3.7	
	Bottom	8.4 ± 7.2	0.05	5.9 ± 4.4	**0.02***	4.9 ± 4.0	0.66
Visioception	Top	6.0 ± 6.6	0.12	4.7 ± 4.3	0.23	5.7 ± 5.6	**0.00****
	Middle	4.7 ± 4.7		3.9 ± 4.5		3.4 ± 3.0	
	Bottom	7.8 ± 6.9	**0.00****	5.3 ± 4.6	0.06	4.6 ± 4.6	0.14

Crossing error means plus or minus one standard deviation (SD) are listed for the first match after the tutorial (training start), the last match of the training session (training end) and the last match of the follow up session. At each time point, we compared z-scores of the crossing errors for the top and middle targets (listed in the “Top” target rows) and the bottom and middle targets (listed in the “Bottom” target rows) using the two-sample t-test. P values are listed for each comparison, e.g. at training start for Video-only there was no significant difference between crossing errors for the top and middle targets ($p = 0.79$). P values for the “Middle” target rows are left blank because we made only two comparisons per time point.
$*p < 0.05$; $**p < 0.01$.

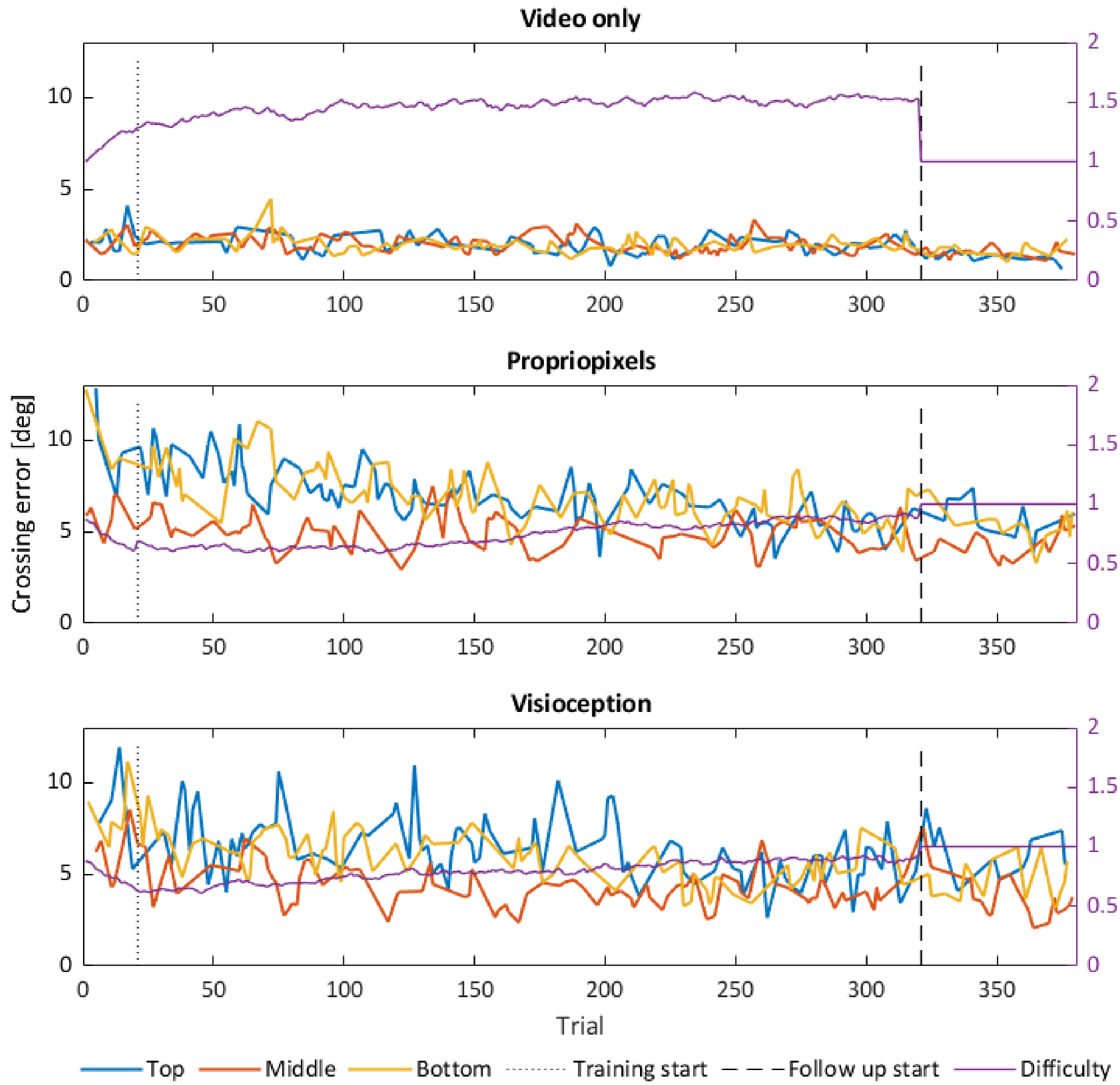

Figure 14. P-Pong ensemble means of crossing error and difficulty per target during play. Crossing errors for each of the three targets (top, middle, bottom) and the difficulty setting are plotted from the start of the last tutorial match (when the success control algorithm was turned on) to the end of the last follow up session match. Each line is an ensemble mean across participants. The start and end of training are indicated by vertical dotted and dashed lines, respectively. During the first session, difficulty was modulated to drive the player to 80% success. When difficulty was ≤ 1, the paddle height was varied and when difficulty was > 1 both the paddle height and the range of vertical ball speeds were varied. Difficulty was initialized at 1, then modulated by the success control algorithm in real time.

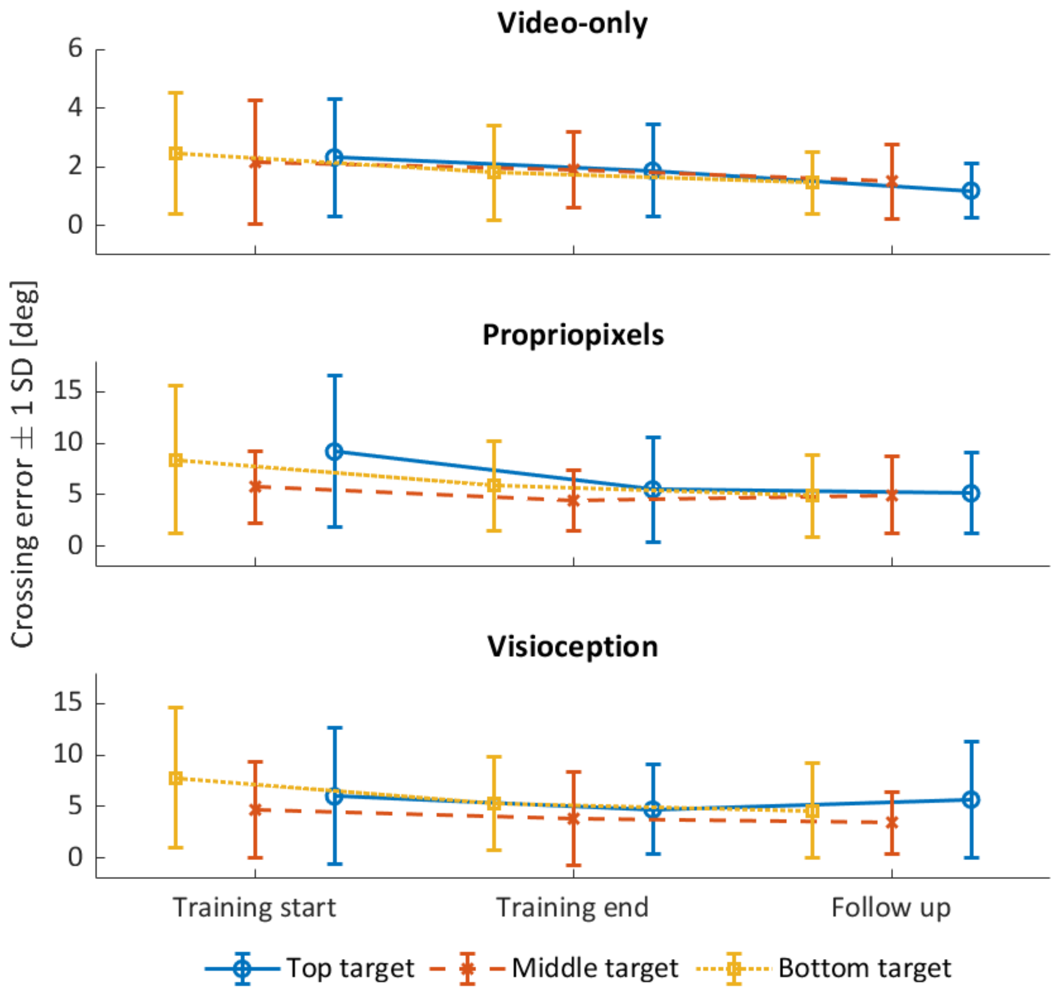

Figure 15. P-Pong crossing error means ± 1 SD at each time point used for hypothesis testing. Training start is the beginning of the first match after the tutorial, training end is the last match of session one, and follow up is the last match at the end of the follow up session. We calculated each mean and standard deviation (SD) as follows. At each time point-target pair, we found the mean of the first five crossing errors per participant and plotted the mean and standard deviation across those participant means. Note that the crossing errors are plotted in degrees, while within subjects z-scores were used for hypothesis testing. The y axis range for the Video-only group (top plot) is less than the range of the Propriopixels and Visioception group plots.

B. Training effect on passive position sense measured by Crisscross

There were small and non-significant variations in Crisscross crossing error across all groups and time points, rejecting our hypothesis that all groups would improve. Means and standard deviations of crossing errors ranged from 4.6 ± 0.8 SD deg to 5.6 ± 1.7 SD deg (Table 11). Both the RM ANOVA model terms, time point, and group-time point interaction, were not significant (p = 0.36, p = 0.34, respectively). Crossing errors are plotted per time point in Figure 16.

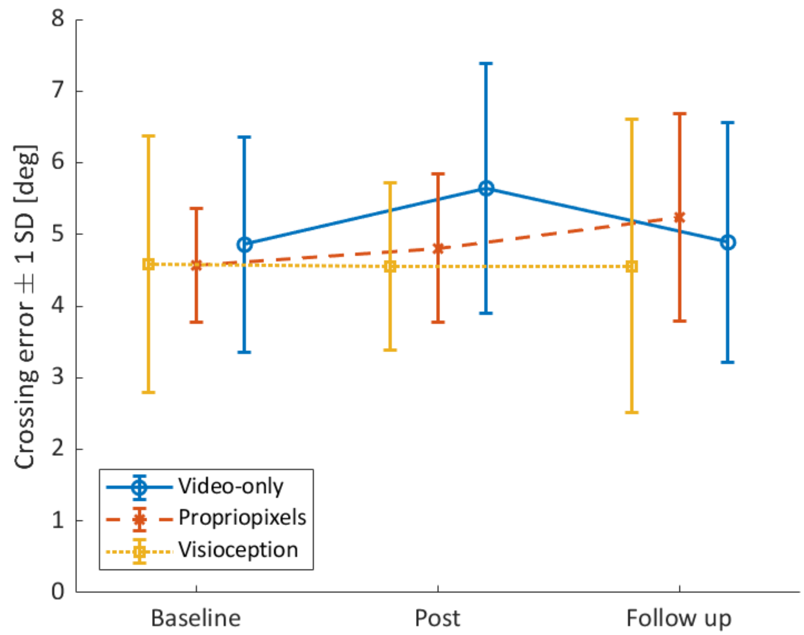

Figure 16. Crisscross crossing errors per time point.

C. *Training effect on dexterity measured by standard and blindfolded BBT*

Scores improved across groups and time points for both the standard and blindfolded BBT tests, rejecting our hypothesis that scores would only improve in the Propriopixels group for the blindfolded BBT. For the standard assessment, baseline scores ranged across groups from 69.1 ± 7.0 SD to 75.7 ± 5.4 SD blocks. All groups increased at post, with mean score increases of 7.0, 3.0, and 5.9 blocks for the Video-only, Propriopixels, and Visioception groups, respectively. The Video-only and Visioception groups did not retain that increase, with follow up scores of 72.8 ± 7.9 and 71.8 ± 6.6 SD blocks, respectively, while the Propriopixels group did with a follow up score of 78.6 ± 6.4 SD blocks. Similarly, for the blindfolded test all groups improved from baseline to post, with scores increasing from 49.3 ± 7.6, 52.3 ± 5.5, and 50.4 ± 3.4 SD blocks to 54.1 ± 7.1, 56.5 ± 5.1, and 56.7 ± 5.8 SD blocks for the Video-only, Propriopixels, and Visioception groups respectively. The mean scores decreased by 1.5 to 2.6 blocks at follow up (Table 11). Scores are plotted across time points for both assessments in Figure 17.

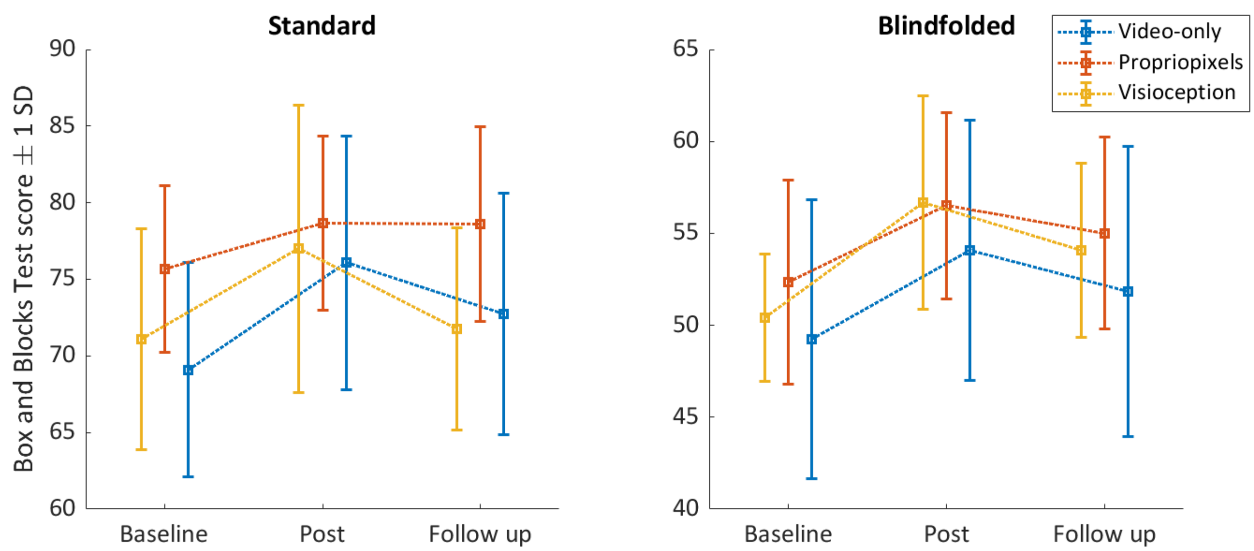

Figure 17. Box and Blocks standard and blindfolded assessment scores at each time point.

We fit an RM ANOVA model with the same two-terms structure to test for time and group-time interaction effects. BBT scores for both assessments were normally distributed and therefore no transformation was required. For both assessments, the time terms were significant ($p < 0.01$), and the group-time interaction term was not significant ($p=0.19$, $p = 0.65$ for standard and blindfolded assessments, respectively). There were significant differences between all the time points ($p < 0.01$ for all three time point combinations), indicating that the observed changes in means, increasing from baseline to post, decreasing from post to follow up, and net increase from baseline to follow up were all significant for both assessments.

D. Success control algorithm performance

With a desired success rate of 80%, the success control algorithm regulated success to 79.1 ± 3.2 SD % at the end of training. Success rate trajectories for each participant are plotted in Figure 18.

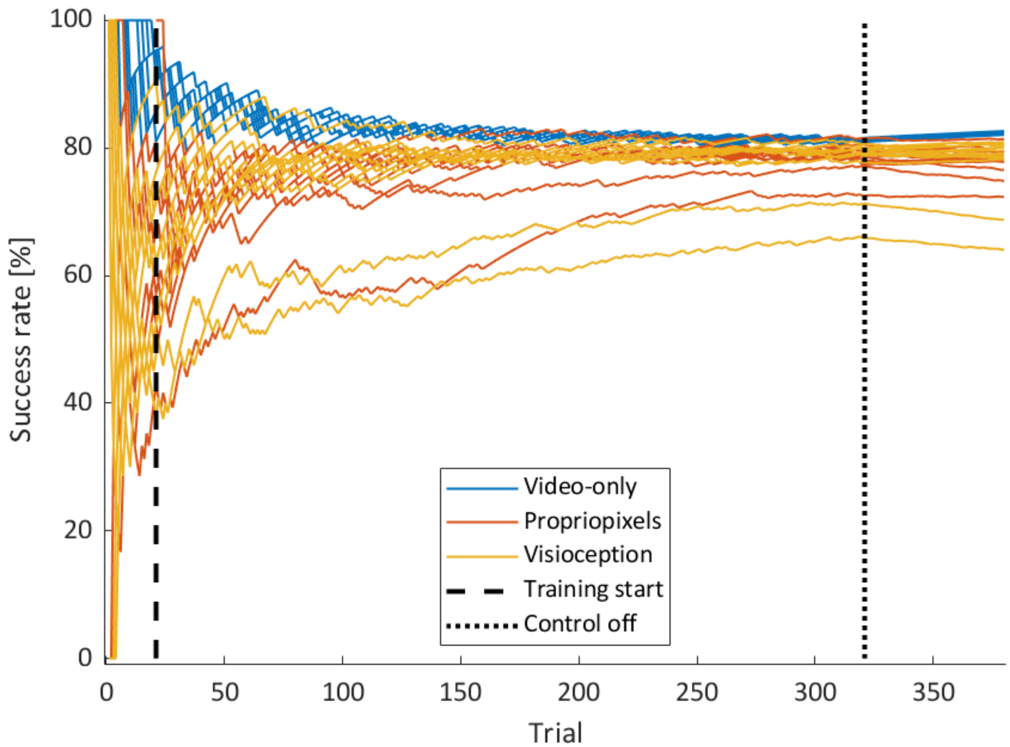

Figure 18. Success rate trajectories for each participant during P-Pong play.

Success rates are plotted per experimental group from the moment the control algorithm was switched on (during the last match of the tutorial) through the end of the follow up session. From the algorithm's perspective there was no difference between the tutorial and training matches, however we delineate between the periods of this session (training start, vertical dashed line) to be consistent with our experiment descriptions. At the end of the training session (vertical dotted line) we turned the algorithm off and reset the difficulty to its initial value (1 on the interval [0, 2]). The success rates diverge from this time point at a slow rate due to the way success rate is calculated – as the number of trials increases, the rate of change of the success rate decreases. Note that the initial success rate (trial one) can take on one of two values: 0% or 100%.

We used regression to verify that the algorithm converged player success to the desired 80% rate. Based on the hockey stick-like shapes of the trajectories, we fit a non-linear model to explain the rate of change of the success rate error ε_s as a function of play time t with the form $\varepsilon_s(t) = c_0\varepsilon_0 e^{c_1 t} + c_2$, having two independent variables initial success rate ε_0 and trial count t , and an intercept term. The model yielded an exponential decay of $\varepsilon_s = 0.73\varepsilon_0 e^{-0.033t} + 3.1$ % ($p = < 0.001$, $R^2 = 42\%$), indicating a model estimated final success rate error of 3.1 (95% CI 3.0-3.3) % after 320 trials (approximately 30 minutes of play). Note that the negative exponential rate (-0.033 95% CI -0.032 – -0.034) and small intercept terms (0.73 95% CI 0.71 – 0.75) are the most meaningful portions of this model – although p-values in time series regression are susceptible to being falsely low due autocorrelation, we did not constrain the terms to specific ranges, e.g. the exponential rate could have taken on a positive or negative value, so it is not the statistical significance that is of importance as much as it is the interpretation of the model coefficients. The success rate error evolution during P-Pong play and model estimated errors are shown in Figure 19.

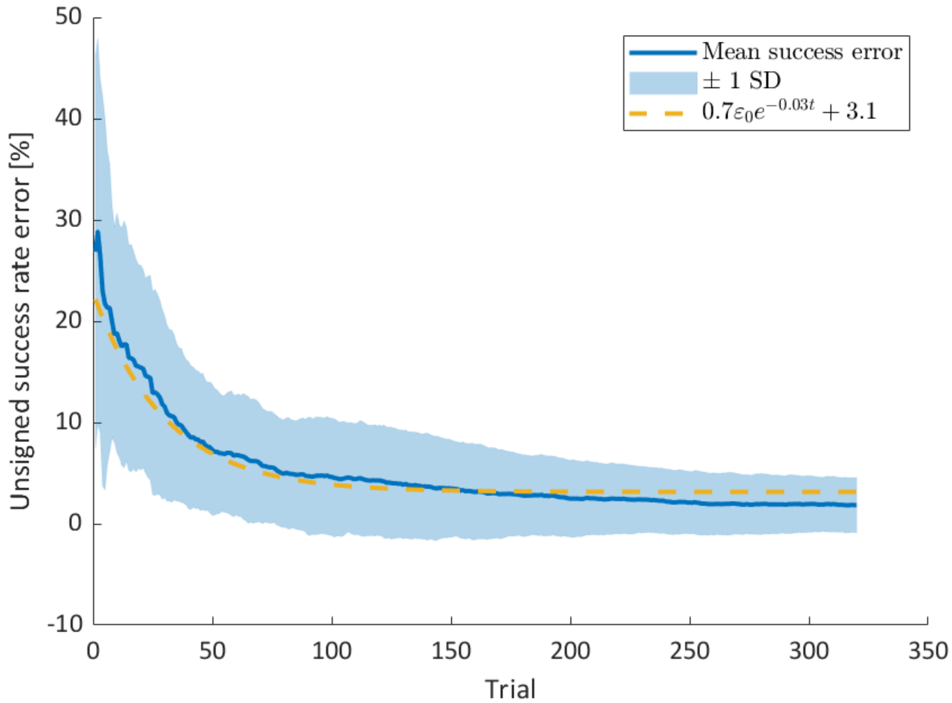

Figure 19. Success rate error trajectory and exponential decay regression model fit.
The ensemble mean of success rate error across all participants at each trial (solid line) ± 1 standard deviation (SD) (shaded region) are plotted with the estimated values from the non-linear regression model (dashed line). The trial t ranges from 1 to 320, therefore when t is small the exponential term dominates, and at the end of the trial it asymptotically approaches zero yielding a final model estimated error of the intercept term 3.1%.

E. Passive proprioception acuity dependence on speed (Crisscross)

We fit a linear Crisscross crossing error versus approach speed model for each participant. Without correcting for multiple comparisons, only three of the 36 linear models had significant slope terms (p = 0.01, p = 0.04, and p = 0.04). After correction all linear models were non-significant. The coefficients, standard errors, p values, and model R^2 values are listed in Table 13 for the models with the 10 lowest unadjusted p values. The linear models for the three participants with significant dependence on approach speed before correction are plotted in Figure 20.

Table 13. Crisscross approach speed effect on crossing error linear regression models.

		Slope [sec]				Intercept [deg]		
	R^2	Estimate	SE	p	p Adjust	Estimate	SE	p
1	0.30	-0.31	0.11	0.01	0.52	12.14	2.60	0.00
2	0.23	0.25	0.11	0.04	0.69	-1.48	2.78	0.60
3	0.23	-0.12	0.05	0.04	0.48	6.97	1.34	0.00
4	0.15	-0.13	0.07	0.10	0.92	7.66	1.70	0.00
5	0.15	-0.19	0.11	0.11	0.76	9.50	2.57	0.00
6	0.14	0.17	0.10	0.12	0.71	0.29	2.91	0.92
7	0.13	-0.06	0.04	0.13	0.67	3.96	0.93	0.00
8	0.12	-0.09	0.06	0.14	0.63	5.42	1.40	0.00
9	0.10	-0.11	0.08	0.18	0.71	6.04	2.00	0.01
10	0.09	-0.07	0.05	0.22	0.78	5.60	1.26	0.00

No models yielded significant p values < 0.05 after adjustment using the false discovery rate method described in [130]. Approach speeds were in units of deg/sec. Abbreviations: SE, Standard error; p Adjust, adjusted p value.

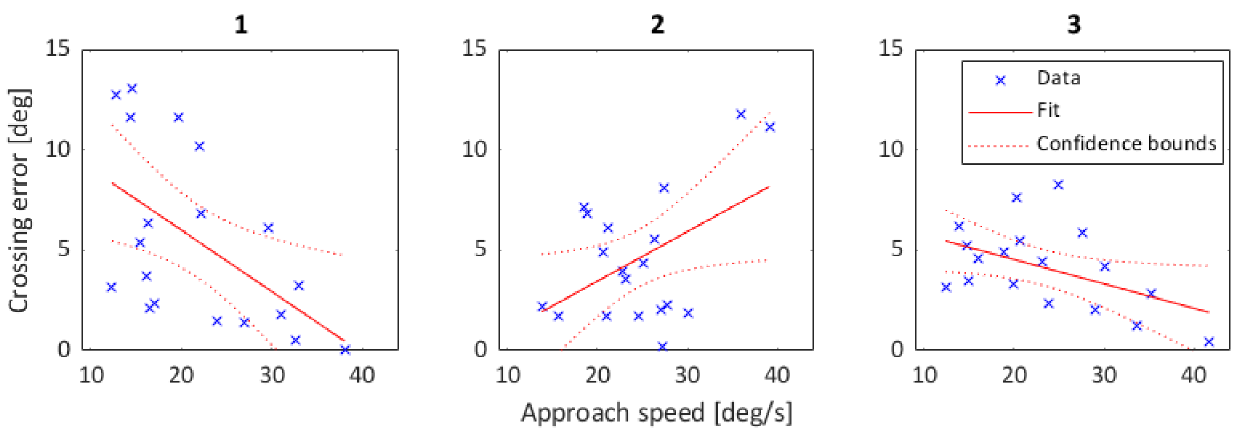

Figure 20. Relationships between Crisscross crossing error and finger approach speed for three lowest unadjusted p values.

Each plot is labeled with the number of the corresponding row in Table 13. Crossing errors decrease as speed increases for 1 (left) and 3 (right), whereas crossing errors increase as speed increases for 2 (middle).

F. Investigation of the specificity of training to velocity

We investigated potential causes of our non-significant Crisscross crossing error analysis results. First, we investigated potential effects of training and testing at different speeds, which we describe in this subsection. Second, we investigated potential Crisscross assessment floor effects, which we describe in the next subsection.

Was there a difference in the P-Pong and Crisscross approach speeds, and if so, were there only Crisscross crossing error improvements in those speeds? First, we compared distributions of P-Pong and Crisscross approach speeds. For P-Pong, approach speed was calculated in the same way as Crisscross, only using the velocities of the ball and paddle. We found that P-Pong approach speeds were lower than Crisscross.

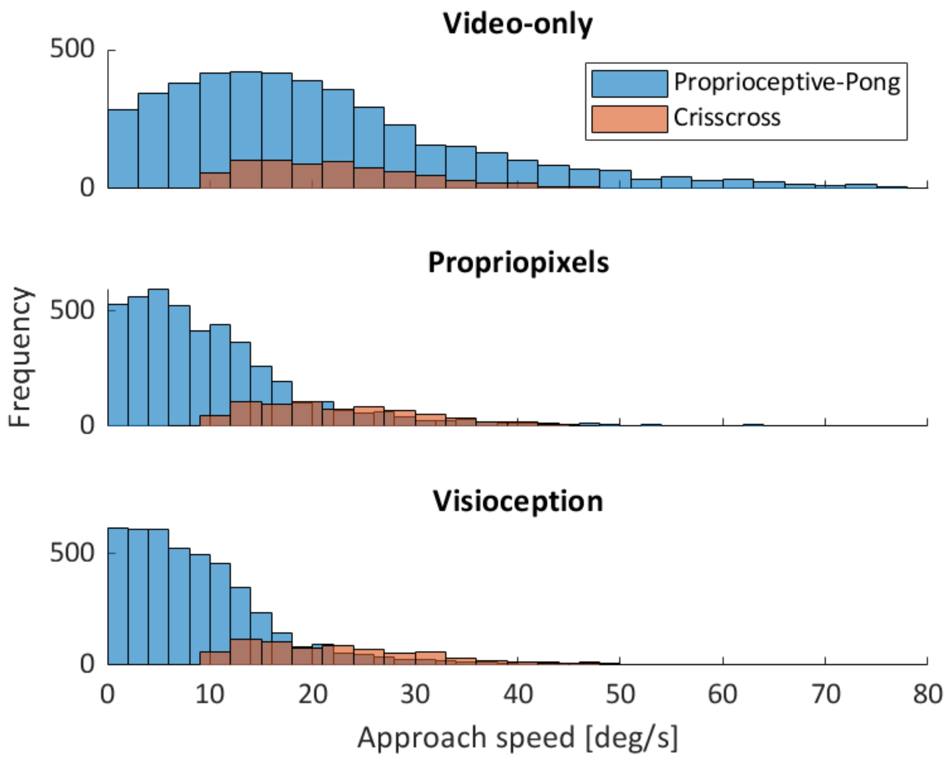

Figure 21. Approach speed comparison between the training (Proprioceptive-Pong) and passive proprioceptive assessment (Crisscross) activities.

We repeated the same RM ANOVA analysis used to evaluate Crisscross training effects with only "slow" trials that fell below the 20th percentile of speeds (the slowest speeds). We again applied a within subjects z-score transform to correct for non-normality in the data. The RM ANOVA result was no different, with non-significant p values for the model terms time (p =

0.24) and group-time interaction (p = 0.55). These results suggest that the difference in approach speeds did not limit the response of Crisscross to P-Pong training.

G. Investigation of Crisscross assessment floor effect

Considering that baseline proprioceptive acuity was high with a mean crossing error of 4.7 ± 3.3 SD deg across all groups, did the Crisscross assessment not adequately challenge passive proprioception acuity such that improvements (that did in fact occur) were not measured? We first investigated whether there were improvements at faster approach speeds and analyzed only Crisscross trials that were greater than the 80th percentile of approach speeds. We performed the same RM ANOVA analysis, first transforming the crossing errors with within-subjects z-scores and then fitting a model of the same structure with time and group-time interaction terms. From a significance perspective the result was no different: p values for the time (p = 0.19) and group-time interaction (p = 0.72) model terms were not significant. Therefore, within the range of speeds tested, there was no difference in improvement at the faster speeds.

To further investigate potential floor effects, we next compared baseline Crisscross crossing errors from this chapter to the previous chapter (we assume there was no consequential floor effect in the previous chapter since a group significantly improved). As we performed Crisscross with different test protocols and devices, we attempted to identify whether any differences had a measurable effect. To make this comparison we mapped the crossing errors measured by PINKIE on a linear scale in millimeters to the same space used in this chapter, a metacarpophalangeal joint angle (MCP) in degrees. We assumed that furling/unfurling MCP trajectories were the same for both robots. We then mapped each participant's errors from millimeters to a percent range of motion using their PINKIE range of motion settings (these were specific to each participant), then mapped that 0-100% value to an MCP angle using the same range of motion to MCP mapping as the FINGER robot. We then compared the baseline Crisscross crossing errors between the current chapter (performed with the FINGER robot) and last chapter (performed with the PINKIE robot) using the Wilcoxon Rank Sum test.

The median FINGER Crisscross crossing error 4.2 (IQR 2.2 - 6.2) deg was significantly lower than the PINKIE crossing errors 8.3 (IQR 3.8–13.8) ($p < 0.01$). The comparison of FINGER and mapped PINKIE crossing errors is shown in Figure 22.

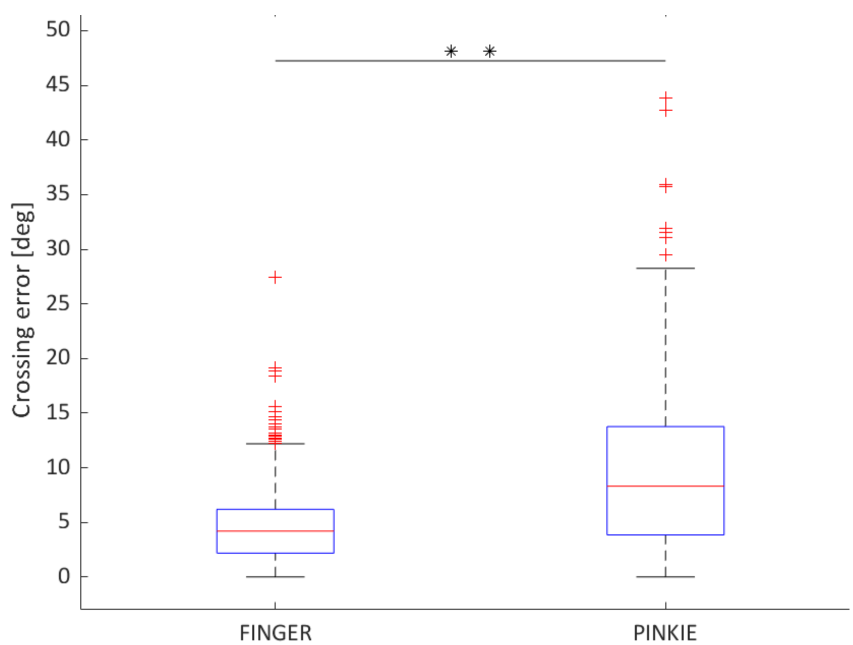

Figure 22. Comparison of baseline Crisscross results between the current and last chapter. PINKIE crossing errors were significantly greater than the FINGER crossing errors with **p < 0.01.

Discussion

Building on our work described in the Chapter 3, we first described updates to the P-Pong game: a non-robotic proprioception-targeting input output/structure that we call "Visioception", a target based success system where instead of simply returning the ball the player’s goal is to hit the ball to a specific location, and a revised success control algorithm designed to both up- and down-regulate success for a range of player. Different from its previous implementations, the success control algorithm controlled only virtual parameters – vertical ball speed and/or paddle height to regulate success, meaning that although we carried out the following experiment with a complex robot capable of physical assistance, it could have been performed on a less complex binary impedance robot like PINKIE. We next evaluated the effects of the proprioception targeting non-robotic and robotic gaming paradigms Visioception and Propriopixels against a more typical Video-only training paradigm. We evaluated their immediate and short-term retention effects on proprioception acuity, which we measured robotically using the Crisscross assessment and in-game P-Pong game performance, and dexterity using a standard and blindfolded version of the BBT using RM ANOVA. Across all outcome measures, there were no significant interaction effects between group and time point, meaning that no one group significantly improved more than the other. P-Pong crossing error, standard BBT, and blindfolded BBT significantly improved over time ($p < 0.01$ for all outcomes), where crossing error significantly decreased for P-Pong and were retained at follow up, and BBT scores improved post training, and had significant decreases post training with overall improvements from baseline to follow up. Crisscross performance however did not significantly change over time ($p = 0.36$), with small <1 deg fluctuations in crossing error across groups and time points. We investigated potential causes of this result and found that Crisscross crossing errors were significantly lower ($p < 0.01$) than the previous chapter at baseline. Next, we evaluated the performance of the success control algorithm and found that it drove participants to the desired 80% success rate with a final rate of 79.1 ± 3.2 SD % across all groups completely virtually -

success was upregulated by increasing paddle height and downregulated by decreasing paddle height and increasing vertical ball speed. We verified its time series convergence using nonlinear regression and found that an exponential decay model explained 42% of the variance in success rate errors over the duration of P-Pong play. Lastly, evaluated the effects of varying approach speed during the Crisscross assessment using linear regression and found that crossing error was not significantly correlated with approach speed after correcting for multiple comparisons. We next discuss these results, investigate potential causes, and propose a set of future directions.

A. Effect of difficulty on P-Pong crossing error

For the Video-only group the initial difficulty was too low while for the other groups it was too high to achieve our 80% desired success rate. This is reflected in the difficulty trajectories: for the Video-only group the mean difficulty during training was higher than the initial difficulty setting (1.0) whereas for the other groups it was lower (Figure 14). When we compare difficulties from the end of training "post" time point to the follow up matches, for Video-only there was a large change in difficulty, whereas for the other groups there is a small change, which correlates with the magnitudes of crossing error changes from baseline to post. Further, for the Video-only group's increases in vertical ball speed and decreases in paddle height may have affected position errors, not just success, as position errors considerably decreased (26%) from post to follow up, i.e. when we manually stepped the difficulty from its high value (~1.5, which was iteratively stepped by the success control algorithm) to 1.0 at the start of the follow up session we observed a crossing error improvement. We make this distinction because theoretically the success control algorithm could have no effect on position error – a player could be too successful because >80% of their position errors fall within the height of the paddle contact area and the algorithm simply decreases that height such that their crossing error distribution maps to an 80% success rate. In this example, crossing error would be independent of difficulty and difficulty would be the mapping from crossing error to success rate. However, the "step" in crossing error observed for the Video-only group implies causality, that increasing paddle height and decreasing approach speed caused a decrease in crossing error (we cannot decouple whether one and/or the other was causal).

B. Training effect on active position sense quantified by P-Pong

P-Pong crossing errors decreased over the course of training for the proprioception-targeting training groups with 32% and 26% decreases in mean crossing error for the Propriopixels and Visioception groups, respectively, indicating that both groups improved active finger position sense. In addition to groups improving overall, there was an initial crossing error dependence on target location at training start(there was a significant difference between top and middle or top and bottom for both groups) that decreased over the course of training. This can be seen in Figure 15 where at training start there were differences in the means and standard deviations of the top, middle and bottom targets and by training end the discrepancies decrease. For the Propriopixels group at follow up there was no significant difference across targets. However for the Visioception group, there was no significant difference across targets at training end, however this was not retained at follow up.

For all groups, changing the target corresponds to changing the virtual center of the paddle contact area. For example, a middle target corresponds to the desired contact area being centered around the index finger whereas a bottom target corresponds to that center being offset “above” the index finger. There were key differences in how that offset was conveyed. In the Visioception mode it was presented visually (the small triangle indicated the center of the contact area), requiring players to integrate visual ball information from the screen, paddle offset information from the screen, and paddle position information from intrinsic finger proprioception. For example, if the ball approached below the paddle then the player needed to move their finger down further for a bottom target than they did a middle target; how much farther they needed to move was conveyed visually, and the amount they moved was sensed proprioceptively. Quite differently in the Propriopixels mode, players were not given any information about how much “extra” they need to move their finger for a bottom vs middle target. Meaning that again if the ball were approaching below their paddle, then for a middle target they would need to align their index with their middle finger, which is intuitive. However, for a bottom target they did not initially know how far to move their index finger below their middle finger – in this sense the anatomical locations of the top and bottom targets were completely novel. And yet, they learned it during training and retained that learning over a 1-3 day period.

Multiple factors could explain the differences in learning the top/bottom target position “offset” from the middle target between the Propriopixels and Visioception groups. The differences could be attributed to increased uncertainty associated with the offset being conveyed through different afferent pathways versus the same pathway, where the Visioception group compared visual and proprioceptive information while the Propriopixels group only compared proprioceptive information. Alternatively, the Visioception group could have reduced top/bottom target crossing errors less because visual information dominated useful intrinsic proprioceptive information, which could be explained by the specificity of learning hypothesis. The specificity of learning hypothesis states that "learning is specific to the source or sources of afferent information that are more likely to ensure optimal task performance" [25]. If we broadly classify the different game modes as types of feedback, then Propriopixels mode may have better guided participants by forcing them to use only proprioceptive information. Visioception instead allowed participants to use both proprioceptive and visual information which has been shown to dominate other afferent information, such as proprioception [24]–[27]. Put another way, the visual feedback information provided during Visioception may have caused participants to ignore useful intrinsic proprioceptive information. Of note, we did not provide any haptic guidance to participants, e.g. in the form of physically assisting them to a desired finger offset distance, which has been shown in several studies to benefit learning [131]–[136], and yet the Propriopixels group learned the target task and retained it. Perhaps "guiding" them to simply focus on intrinsic proprioceptive information as opposed to visual information caused them to learn and retain the target task.

C. Training effect on passive position sense measured by Crisscross

Crisscross crossing error did not improve in any groups across any of the time points. We found this result surprising and interesting in light of our results in Chapter 3 where crossing error significantly improved for the Propriopixels group, especially considering that the Propriopixels group performed 50% more repetitions, trained for twice the time, and achieved considerably higher success rates in this experiment. The study cohorts were similarly distributed in age, sex, and handedness, and all participants were unimpaired and recruited from the same geographical area. Further, all groups improved in an active movement tracking task as crossing error significantly decreased during P-Pong training for all groups, raising the question: why did Crisscross crossing error not respond to P-Pong training? We identified potential differences that could have caused the lack of Crisscross training response stemming from changes to the P-Pong game and Crisscross assessment. We investigated the potential cause of specificity of training to velocity by analyzing whether participants improved in Crisscross only at approach speeds that were trained by P-Pong and found no significant effect. The P-Pong speeds were slower than Crisscross for the Propriopixels and Visioception groups and there was no difference in Crisscross improvement at slower speeds.

D. Evidence of Crisscross assessment floor effect

We investigated the potential cause of Crisscross assessment floor effect in two ways. First, we analyzed training effects at only the fast Crisscross speeds, however we saw no training effect at the fast speeds. Second, we compared baseline Crisscross crossing errors with those measured in the previous chapter using PINKIE and found that crossing errors were significantly lower when measured with FINGER. And although our method of mapping the PINKIE crossing errors into the same MCP joint angle space as the FINGER crossing errors was inexact due to differences in MCP trajectories between the two robots, the difference between the two median errors was large (4.2 deg for FINGER and 8.3 deg for PINKIE). Considering the similarities between the demographics of the two study cohorts, this indicates that the lack of response of Crisscross to P-Pong training could be due to a floor effect. Key differences between the assessments are that the PINKIE version varied the crossing positions throughout the ranges of motion of the index and middle fingers while in FINGER Crisscross they were at the center of ranges of motion, the speed of each finger differed (one finger could move quickly and the other slowly), and the highest speeds were faster than FINGER Crisscross. Ingemanson et al. also measured higher mean crossing errors (6.0 ± 0.4 SD versus 4.7 ± 2.2 SD deg in this chapter) using a similar Crisscross assessment design with the FINGER robot, where the crossing positions and speed per finger were varied [106]. Further, in this study there was no significant correlation between crossing error and approach speed, which may indicate that the fastest tested speeds were too slow. Based on this result, the lack of improvement in Crisscross after P-Pong training could be attributed to an assessment floor effect.

E. Evidence of specificity of training to movement type

Considering that there were significant improvements in P-Pong crossing error, a measure of active position sense for the proprioception-targeting training groups, and no significant improvements in Crisscross crossing error, a measure of passive proprioception, the lack of Crisscross improvement could be caused by a specificity of training to movement type. We changed the P-Pong control scheme from passive to active movement. In the previous chapter the index paddle finger was passively moved – participants relaxed their index finger and controlled it using a thumb stick held in their contralateral hand. In this chapter participants controlled the paddle by actively moving their index finger. This difference could explain why in the former chapter Crisscross crossing error responded to P-Pong training, whereas in this chapter it did not. Our results from this chapter align with those reported by Winter et al., that active movement training tended to cause motor and proprioceptive gains [121]. However, in contrast to our results, while most studies that performed active movement training tended to quantify proprioception with active joint position sense error, some did report increases in passive joint position sense error. These studies did administer much higher doses over a 6–12-week training period, therefore the lack of Crisscross response to active movement training observed in this study could in fact be a reduced response. Next, we provide a set of recommendations for future directions.

F. Limitations and future directions

The lack of a "sham training" control group limits our ability to interpret the BBT results. As all groups improved, it is unclear whether that improvement is caused by P-Pong training or simply repeating the BBT. To our knowledge, the test-retest reliability of BBT scores has not been established for unimpaired participants. In a recent systematic review, Winter et al. found that active movement training was most successful in improving proprioception and motor capacity [121]. Therefore, it is feasible that BBT in all groups improved due to repeating the test.

Considering the potential assessment floor effect with version of Crisscross implemented with FINGER, in future work we suggest increasing the difficulty of the assessment. This could be accomplished by implementing the elements of the PINKIE Crisscross assessment that were potentially more proprioceptively challenging: varying the position where the fingers cross, moving each finger at an independent speed, and increasing the fastest tested speed.

Perhaps one of the most interesting directions from this study is the potential specificity of proprioceptive learning to active versus passive movement. All groups significantly improved their P-Pong crossing errors while no groups significantly improved their Crisscross crossing errors. Lacking an assessment that is sensitive to P-Pong training effects, it is difficult to compare the benefits of the different active movement P-Pong game modes. Of note, there were clear differences in how the groups responded to the novel P-

Pong targets: they had no measured effect on the Video-only group, while they increased crossing error in the other groups – an effect that decreased with training. The Propriopixels group exhibited excellent learning effects, with minimal differences in error across targets at follow up, which raises the questions: did incorporating targets enhance learning? We believe this is an interesting avenue to explore: the extent to which novel stimuli can enhance propriomotor learning, and through what modalities the stimuli should be conveyed. And given that this learning was achieved without physical assistance, that their benefits could be realized with simple robotic devices like a binary impedance robot.

Considering the potential for specificity of training to movement type, in future work we suggest including proprioception assessments that test both active and passive movement. Additionally, assessments could be designed into training activities where the same block of assessment trials are automatically administered to each group during of game play. If we regard error and success feedback information as forms of guidance that help train participants [137], removing such information could enhance an in-game assessment. For P-Pong this could be implemented by exposing players to an equal number of top, middle, and bottom targets in two display modes, Video-only and Propriopixels, at a prespecified interval. The ball and paddle would disappear when the ball reaches the paddle to reduce position error feedback and remain hidden until the ball is reflected off the opposite wall to reduce success feedback. This block of trials could fix the ball speeds, or step through a discrete set of speeds to assess performance at different difficulty levels. Perhaps most importantly, this same block of trials would be administered to all groups so their performance could be easily compared.

Participants responded positively to the levels and high scores. Many commented on leveling up, a new level being harder (although it was not and the participant was just as successful overall), and perhaps most frequently, talked about beating other (fake) players' high scores. This small game design element added depth that we believe the participants enjoyed. Perhaps it enhanced their attention, motivation and therefore their learning. Keeping players engaged was a serious consideration throughout our design and development process. It was challenging, especially for engineers like ourselves with no formal game design training. We found the participants' reactions to the levels and high scores to be a rewarding and interesting success for us as game designers and a useful tool for future work.

Conclusion

We found that 300 repetitions (approximately 30 minutes) of gamified P-Pong training significantly improved crossing error during play, and that such improvements were retained over a one-to-three-day period. All training groups also improved in standard and blindfolded versions of the BBT. Interestingly, those improvements did not transfer to passive proprioceptive acuity, which based on our investigation may be attributed to an assessment floor effect or a specificity of proprioceptive learning to training movement type, giving insights into how the human sensorimotor system learns, how training should be designed to enhance learning, how assessments should be designed to test learning, and avenues for future investigation. As a secondary result, we demonstrated that player success can be controlled completely virtually, without physically assisting participants.

CHAPTER 5: THE FEASIBILITY OF GAMIFIED MULTIMODAL HAND MOVEMENT TRAINING WITH PROPRIOCEPTIVE-PONG AFTER STROKE

Summary of the Chapter

Proprioceptive-Pong requires players to fuse multimodal sensory information, make cognitive gameplay decisions, and act on those decisions with corresponding finger movements. Given the complexity of this task, and range of cognitive, visual, and sensorimotor impairments in its target user population, we evaluated the feasibility of Proprioceptive-Pong for intense hand movement practice with three chronic stroke survivors. During a single session, each participant played a unique progression of Proprioceptive-Pong modes. We varied the display mode (Video-only or Propriopixels), assistance type (none, physical, or virtual), and targets (on or off). For the two participants that played Propriopixels, we evaluated task comprehension by comparing initial finger (position) crossing errors of their first and last matches, where in the first match they played without targets and in the last match they played with targets – the more difficult condition. The first match mean crossing errors significantly decreased from 11.5 ± 7.2 SD to 6.4 ± 4.5 SD deg ($p < 0.01$) during the last match. Success was also regulated during play using both physical and virtual assistance, with final success rate errors of 3.8% to 18.9% over a relatively small number of repetitions for algorithm convergence (80 – 180 repetitions). Our results indicate that in its current implementation, the success control algorithm is capable of regulating success with both physical and virtual assistance. Importantly, we found that stroke survivors were able to understand and play the most cognitively challenging configuration of the game, demonstrating that gamified multimodal training with Proprioceptive-Pong is feasible for stroke rehabilitation. Following, we evaluated the effects of intense, multi-session Propriopixels training on upper extremity propriomotor capacity

and performance with a single stroke survivor. We found that Propriopixels training significantly improved passive proprioceptive acuity ($p < 0.01$) and caused clinically meaningful increases in Motor Activity Log Amount of Use (1.7) and Quality of Movement (1.6) subscales, indicating that Propriopixels training can improve arm use after stroke.

INTRODUCTION

In the previous chapter we evaluated the effectiveness of robotic and non-robotic proprioception targeting gaming paradigms in unimpaired participants, toward the goal of investigating their training benefits in stroke survivors. In this chapter, we present a proof-of-concept study to evaluate the feasibility of playing Proprioceptive-Pong for intense hand movement practice. We find this step crucial given the game's complexity: it requires players to fuse multimodal sensory information, make cognitive gameplay decisions, and act on those decisions with corresponding finger movements.

This multimodal fusion of afferent information we refer to that occurs during Propriopixels has key differences from previous "multimodal feedback" studies. We define multimodal feedback as providing task information concurrently through multiple afferents e.g. visual and haptic guidance during a steering task [138]. Sigrist et al. reviewed multimodal feedback and its multiple potential benefits such as optimized neural activation and neural representation, reduction of cognitive load due to distribution of information processing, potential to complement specificity-of-learning by optimizing feedback information communicated to each modality [139]. In Propriopixels with targets however, while the player is making decisions we split feedback between vision and proprioception. This raises the question, if we do not provide redundant multimodal information that a user can theoretically choose to use or ignore and instead communicate information through different afferents that they must compare (to be successful), then can we attain similar learning benefits? And perhaps more importantly, will this impact task comprehension and ultimately impede learning? With these questions in mind, we next investigated the feasibility of P-Pong as a hand training activity for stroke survivors.

The primary of aim of this chapter was to evaluate the feasibility of improving hand propriomotor capacity and performance after stroke using P-Pong. Specifically, we trialed a progression of game settings to build participants up to the most complex game configuration (Propriopixels with targets) in a single familiarization session. One stroke survivor subsequently performed Propriopixels training over three weeks. We also implemented physical assistance - an application of the same success control algorithm we presented previously and evaluated its performance, which we next present.

METHODS

A. Success control algorithm implementation

We added physical assistance to the P-Pong game and FINGER robot using the same success control algorithm presented in Equation 3. Previously we assisted players by modulating paddle height. With physical assistance, the algorithm instead varied Proportional-Derivative control gains to track assistance trajectories. Each time the player triggered assistance by exceeding a force threshold of >= 3N measured by load cells located at the proximal and mid phalanxes, the FINGER robot tracked a minimum jerk trajectory to the ball-paddle contact location. When assistance was minimized by the algorithm (difficulty > 1) the control gains were zero, yielding no tracking, and when they were maximized (difficulty = 0), the control gains were high and the robot provided high assistance forces to the “successful” paddle location. The success control algorithm was again set to an 80% desired success rate. We made no other updates to the P-Pong game.

B. Participants and experimental design

This study was approved by the University of California, Irvine Institutional Review Board and participants provided consent. Three stroke survivors participated in this study. The inclusion criteria were ≥ 18 years of old, experienced a single stroke ≥ 6 months previously, and ability to score ≥ 3 on the Box and Blocks Test. The exclusion criteria were severe aphasia (score of 3 on the National Institute of Health Stroke Scale), coexistent major neurological disease, and coexistent major psychiatric disease. All participants were male. Additional demographic and stroke information are listed in Table 14.

Each participant first completed four baseline assessments, the Fugl-Meyer Assessment for Upper Extremity, Box and Blocks Test, National Institute of Health Stroke Scale, and Crisscross assessment, then played through a unique set of P-Pong familiarization matches. All clinical assessments were performed by a licensed physical therapist. We performed the Crisscross assessment as we described in the previous chapter with 20 crossing trials and 10 different speeds. The Crisscross assessment outcome was calculated as mean crossing error (the mean of the unsigned difference in finger positions). All baseline assessments results are listed in Table 14.

Table 14. Participant demographic information, stroke information, and baseline assessment results.

	Age	Dom hand	Affected arm	Days after stroke	Stroke type	FMAUE	BBT	NIHSS	Crisscross [deg]
P1	65	Left	Right	1566	Ischemic	49	20	2	8.1
P2	67	Right	Right	2398	Hemorr	54	49	1	9.4
P3	42	Ambi	Right	3672	Hemorr	51	34	1	6.2

Abbreviations: P1, P2, P3, participants 1, 2, and 3; Dom hand, dominant hand; Ambi, ambidextrous; Affected arm, upper extremity side more affected by stroke; Hemorr, hemorrhagic; FMAUE, Fugl-Meyer Assessment for Upper Extremity; BBT, Box and Blocks Test; NIHSS, National Institute of Health Stroke Scale.

Each participant played through a different progression of game settings. Across all participants, we varied the display mode (Video-only or Propriopixels), assistance type (none, physical, or virtual), and targets (on or off). Targets being "off" corresponds to classic Pong game play and is analogous to only using the middle target (see Figure 11 for additional details). For one participant, we reduced the amount of finger extension required to move the paddle up to the highest position at the top of the field/screen. All other participants played with the full range of motion of each finger mechanism. We refer to each combination of game settings as a "condition". Each time we changed the condition, we first verbally explained the changes to the participant then coached them as they played. Participants one experienced three conditions, participant two experienced four conditions, and participant three experienced four conditions. Please see Figures 11-12 for a detailed explanation of the P-Pong game and its settings.

As a first step in evaluating P-Pong feasibility, the first participant played the Video-only mode. They played with targets and assistance off for one match (first condition) then played seven matches with targets and physical assistance switched on (second condition). Lastly, we reinitialized the game difficulty and they played four matches with virtual assistance on in Video-only mode with targets on (third condition).

We evaluated our Propriopixels familiarization procedure with participant two. They first played Propriopixels mode with targets and assistance off for three matches (first condition), then they played seven additional matches with targets and physical assistance on (second condition). Lastly, we reinitialized the game difficulty and they played three matches with assistance off and targets on, also in Propriopixels mode (third condition).

We evaluated an updated Propriopixels familiarization procedure with participant three. Instead of beginning with Propriopixels, they started with one match of Video-only mode with targets off and physical assistance on (first condition). We then introduced Propriopixels and targets incrementally, with four matches of Propriopixels (second condition) followed by three matches of targets (third condition). For the Video-only match and first two matches of the following two conditions (5 matches total), the horizontal ball speed linearly ramped from half its full speed (0.13 returns/ sec, or 8 sec/return) to its full speed over the 20 trials within each match. Lastly, we reinitialized the game difficulty and they played three matches with assistance off and targets on, also in Propriopixels mode (fourth condition).

We evaluated the effects of intense Propriopixels training over multiple sessions with participant one. Participant one trained for a total of nine hours over nine subsequent sessions (1 hour per session x 3 sessions per week x 3 weeks). We administered the 30 item Motor Activity Log, the Box and Blocks Test, and Crisscross before and after training. We repeated the Crisscross assessment five times for a total of 100 crossing trials at each time point. The participant was also given feedback during the assessment of their percent crossing error (a value ranging from 0% for no crossing error to 100% for maximum crossing error) after their guess on each crossing attempt.

C. Data analysis

We calculated P-Pong crossing errors in the same way as the previous chapter. We calculated crossing error as the unsigned difference between the index and middle fingers – the metacarpophalangeal joint angle in degrees. We calculated success rate errors as the success rate (total number of ball hits divided by the total number of attempts) minus the desired success rate 80% (positive errors indicate success > 80%). To evaluate Propriopixels comprehension, we compared crossing errors of the first and last Propriopixels matches for participants two and three using the two-sample t-test. To evaluate the effect of three weeks of Propriopixels training on Crisscross crossing error (participant one), we compared baseline and post training crossing errors using the two-sample t-test. We verified the crossing errors were normally distributed using the Anderson Darling test. We performed all data analysis in MATLAB R2021a.

RESULTS

A. Familiarization effects on P-Pong crossing errors

All participants verbally acknowledged understanding the game before beginning play, and before experiencing a new condition. Crossing errors tended to decrease within each condition, e.g. each time we introduced targets crossing errors increased (going from 5.6 ± 4.8 SD deg to 7.4 ± 5.9 SD deg for participant two) then decreased as each player practiced (decreasing to 5.3 ± 3.7 SD deg for participant two). The crossing errors are listed for all participants in Table 15. The evolution of crossing errors with the progression of game conditions played are plotted individually per participant in Figure 24 - Figure 26.

For the participants who played Propriopixels, we compared their first and last matches of Propriopixels play. Compared to the first match, the last matches were more challenging because they were played with targets. Participants two and three significantly decreased crossing error from the first to the last match ($p = 0.005$ and $p = 0.027$, respectively). Participant two decreased mean crossing error from 12.2 ± 8.0 SD to 6.1 ± 4.1 SD deg and participant three decreased from 10.8 ± 6.5 SD to 6.6± 4.9 SD deg. The crossing errors used for hypothesis testing are depicted in Figure 23.

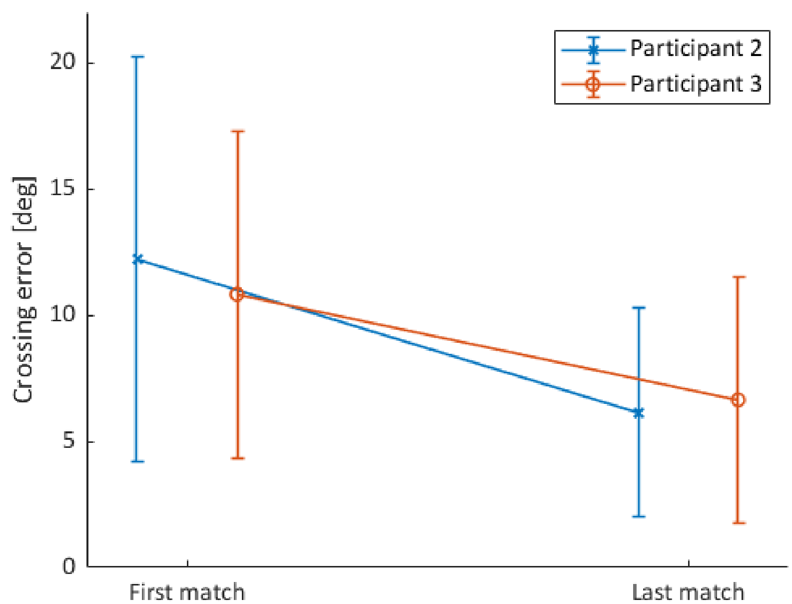

Figure 23. Crossing error comparison for first and last match of Propriopixels play. Crossing errors for all participants that played Propriopixels (participants two and three) significantly decreased from the first to the last match with $p = 0.005$ and $p = 0.027$ for participants two and three, respectively. After playing 260 and 220 Propriopixels repetitions, participant two and three's mean crossing errors decreased to 6.1 and 6.6 deg, respectively. In the previous chapter with unimpaired participants, at follow up the conditions were the same as the "Last match" depicted here: Propriopixels, targets, and success control off with constant difficulty = 1 (no assistance). Participants played 380 Propriopixels repetitions total and attained a mean "Last match" crossing error of 5.0 deg.

B. *Success control algorithm performance*

Success rate errors decreased during play, with mean final success rate errors of 10.7% (participant one, physical assist, 140 trials), 3.8% (participant one, virtual assist, 80 trials), 18.9% (participant two, physical assist, 140 trials), and 9.4% (participant three, physical assist, 180 trials). The success rate errors are individually listed per participant over multiple time points in Table 16.

The evolution of success rate error and difficulty over game play are plotted individually per participant in Figure 24 - Figure 26. Success rate errors are plotted together for all conditions when the success control algorithm was on in Figure 27.

Table 15. P-Pong crossing errors for each participant at different time points.

Event		Participant crossing error mean ± 1 SD [deg]			Success rate error [%]		
		1 (Video-only)	2 (Propriopixels)	3 (Propriopixels)	1	2	3
Start of no targets		10.1 ± 8.6	12.2 ± 8.0	10.8 ± 6.5	32.8	18.1	10.8
End of no targets			5.6 ± 4.8	6.5 ± 4.5		21.0	1.2
Start of targets on	Overall	8.1 ± 5.9	7.4 ± 5.9	9.2 ± 6.7	25.7	18.6	2.4
	Top	9.3 ± 6.4	10.0 ± 5.0	12.1 ± 7.5			
	Middle	8.9 ± 7.3	2.9 ± 3.7	7.6 ± 7.4			
	Bottom	5.8 ± 2.5	8.3 ± 6.8	7.7 ± 4.5			
End of targets on	Overall	2.3 ± 2.0*	5.3 ± 3.7	4.4 ± 2.6	10.7	18.9	9.4
	Top	3.6 ± 3.5*	5.7 ± 3.4	5.6 ± 3.7			
	Middle	1.9 ± 1.2*	4.9 ± 4.7	3.6 ± 1.7			
	Bottom	2.0 ± 1.6*	5.0 ± 3.9	4.5 ± 2.8			
Start of assistance off or virtual assistance on†	Overall	3.3 ± 3.9	7.9 ± 5.3	7.0 ± 4.8	9.2	20.8	9.6
	Top	4.4 ± 6.4	10.7 ± 4.9	8.4 ± 5.2			
	Middle	3.4 ± 1.9	5.4 ± 4.0	7.3 ± 5.2			
	Bottom	2.2 ± 0.9	4.7 ± 4.9	3.9 ± 1.8			
End of play	Overall	8.0 ± 8.0	6.1 ± 4.1	6.6 ± 4.9	3.8	25.7	12.8
	Top	13.6 ± 10.3	6.9 ± 3.6	7.7 ± 6			
	Middle	6.9 ± 4.6	4.5 ± 4.1	6.9 ± 3.5			
	Bottom	2.7 ± 2.8	6.5 ± 4.9	5.6 ± 4.9			

Each event is one match (20 trials), e.g. "Start of no targets" is the first match with no targets – for all participants this was the first match. Participant one played only one no targets match, therefore the "end of no targets" event is left blank. Like the crossing errors, each success rate error column is for an individual participant. Success rate errors are means over the 20 trials within each event.

*Extension range of motion was reduced for participant one between start and end of targets on events.

†The setting changes differed across participants. For participant one, the assistance type changed from physical to virtual. For participants two and three, the assistance was turned off. The difficulty was reinitialized for all participants at this time point.

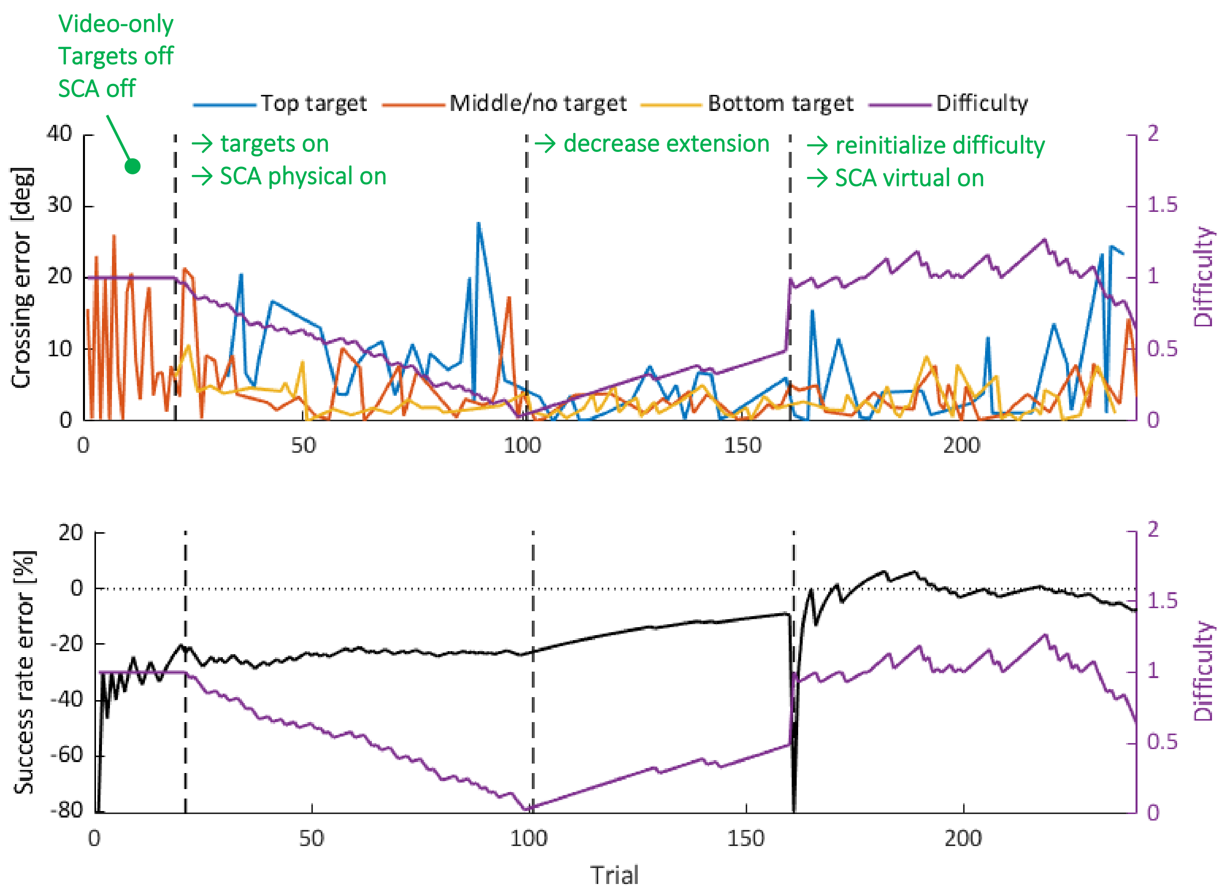

Figure 24. Crossing and success rate errors for participant one.

The first participant played all matches in Video-only mode. They played the first match (20 trials) with targets and the success control algorithm (SCA) off and seven matches (140 trials) with targets and the SCA on using physical assistance . Due to their inability to extend their index paddle finger, indicated in the plot by high top target crossing errors (blue line), we reduced their full extension position (an in-game setting) after 100 trials. Lastly, we reinitialized the game difficulty and the participant played with the SCA on using virtual assistance for four matches (80 trials). The transition between these settings is marked as vertical dashed lines and with settings labeled in green. As assistance increases (lower difficulty) position errors decrease for all targets, accept trial 80-100 when the participant grew increasingly frustrated with their inability to extend their index finger. After reducing their range of motion (ROM), the participant's performance improved and the game difficulty correspondingly increased, decreasing the level of physical assistance. During virtual assist play, the difficulty initially increases, then as position errors correspondingly increase and success rate decreases, the difficulty decreases at the end of play. To more clearly depict virtual assistance performance, we reinitialized the success rate error calculation (reset the hit and miss counts) at 161 trials when we reinitialized the difficulty and changed the assistance type.

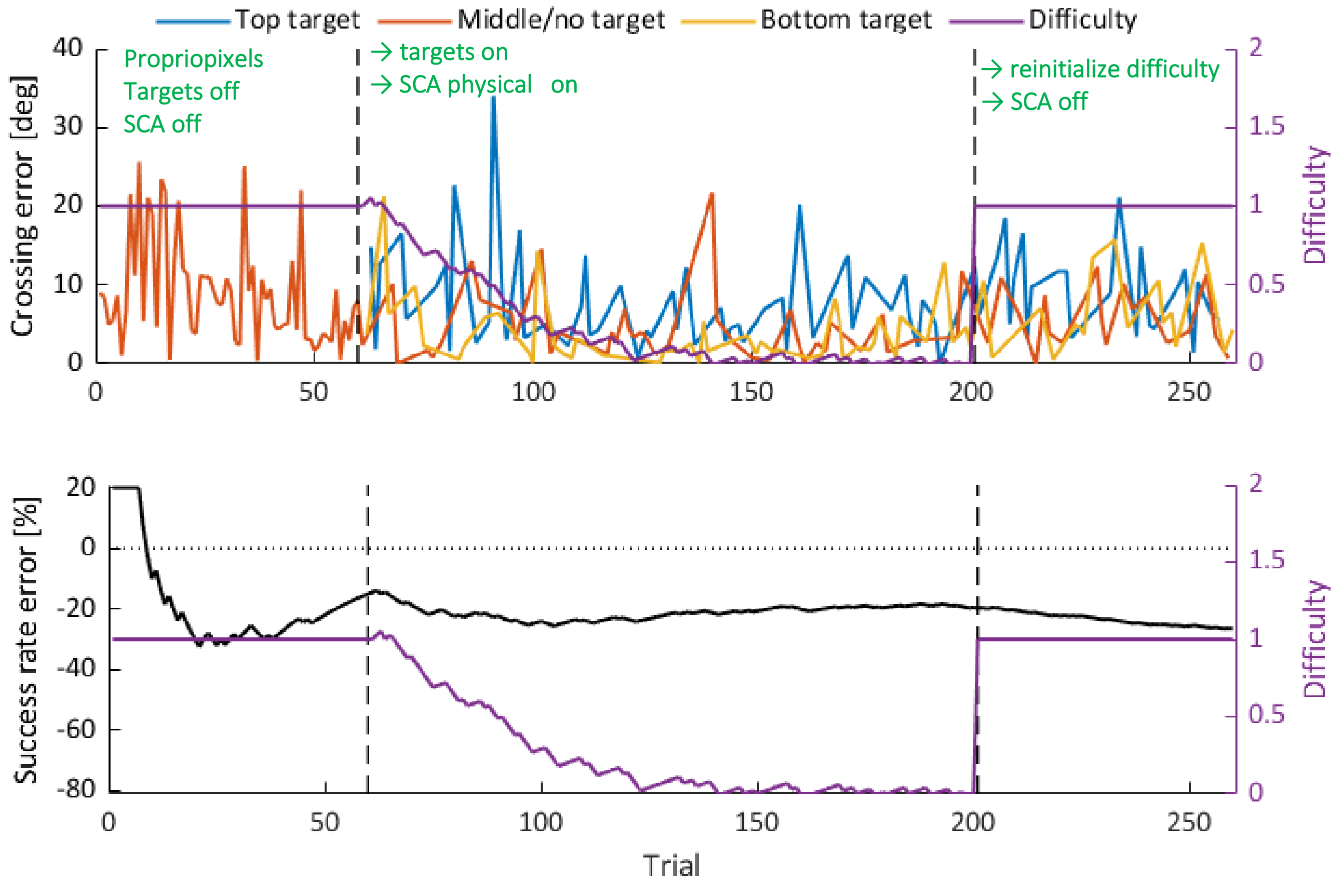

Figure 25. Crossing and success rate errors for participant two.

The second participant played all matches in Propriopixels mode. They played with the full finger flexion-extension range of motion of the robot for all of play. They first played without targets and with the success control algorithm (SCA) off for three matches (60 trials) after which they played with targets and the SCA on using physical assistance for seven matches (140 trials). Following, we switched the SCA off and reinitialized the difficulty. The transition between these settings is marked as vertical dashed lines and with settings labeled in green. Increasing assistance decreased crossing error, but at maximum assistance success rate error was not minimized - 19.3% at 200 trials before we turned assistance off and reinitialized the difficulty. With each setting change, crossing error initially increased, then decreased with practice.

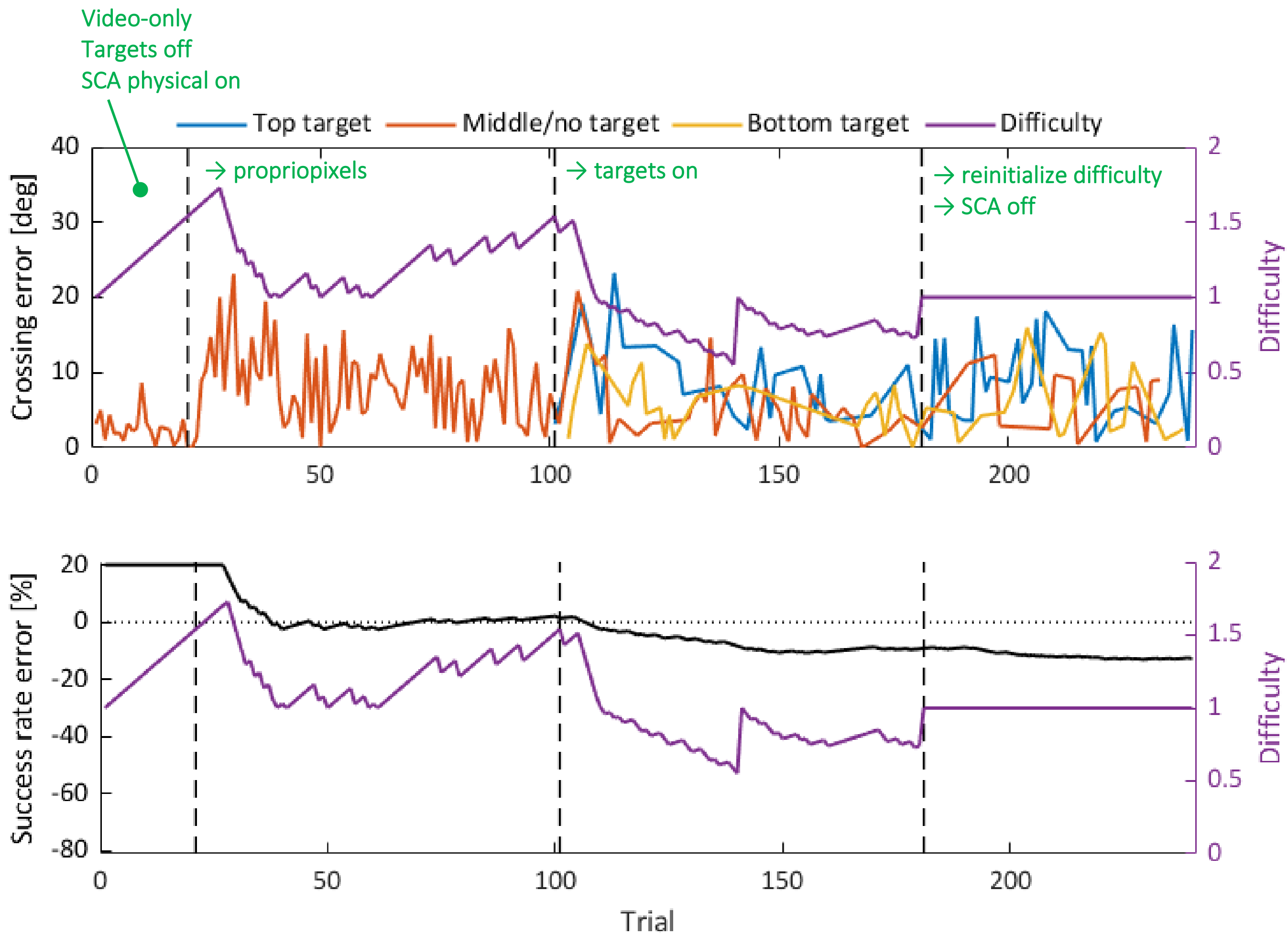

Figure 26. Crossing and success rate errors for participant three.

The third participant played one match in Video-only mode and the remaining matches in Propriopixels mode. They began with the success control algorithm (SCA) on using physical assistance and targets off. After one match (20 trials) we changed the display mode to Propriopixels (no targets) for four matches (80 trials). Next, we introduced targets, which they played for four matches (80 trials). Lastly, we reinitialized the difficulty and switched the SCA off for the final three matches (60 trials). The participant was 100% successful for the first match (Video-only), indicated by the constant 20% success rate error and monotonically increasing difficulty. For difficulties >= 1.0 there was no assistance – the paddle height decreased and the vertical ball speed increased. With each setting change (introduction of Propriopixels, targets, then assistance off) error initially increased, then decreased with practice.

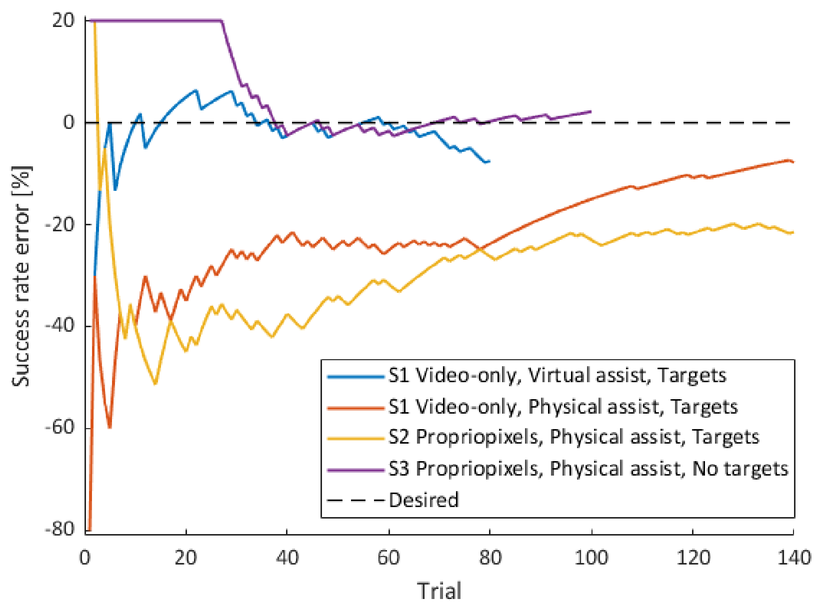

Figure 27. Success rate errors during P-Pong play.

The success rate errors are shown for each participant while the success control algorithm is on with conditions listed in the legend – either Video-only or Propriopixels modes, virtual or physical assistance, and targets or no targets. Participant one is listed twice because their difficulty was reinitialized when they started the virtual assistance mode (after playing the physical assistance mode). To depict success control algorithm performance more clearly, we set the hit and miss counts to zero at the start of each success rate trajectory then recalculated success for participants one and two. This was not necessary for participant three because they began playing with the algorithm on.

C. Effects of multi-session Propriopixels training

Mean Crisscross crossing error significantly improved from 9.8 ± 6.4 SD deg at baseline to 6.4 ± 4.8 SD deg post training ($p < 0.01$) (Figure 28). The Motor Activity Log scores improved from baseline to post training, from 1.4 to 3.1 for the Amount of Use subscale and from 1.5 to 3.1 for the Quality of Movement subscale, respectively. Both improvements are greater than the minimum clinically meaningful changes of 1.0-1.1 reported in [77]. Furthermore, compared to the improvements in amount of use studied in Chapter 2, the participant's increase in amount of use was greater than 93% of the (104) chronic stroke participants.

There were small changes in Box and Blocks Test scores from 21 at baseline to 23 post training.

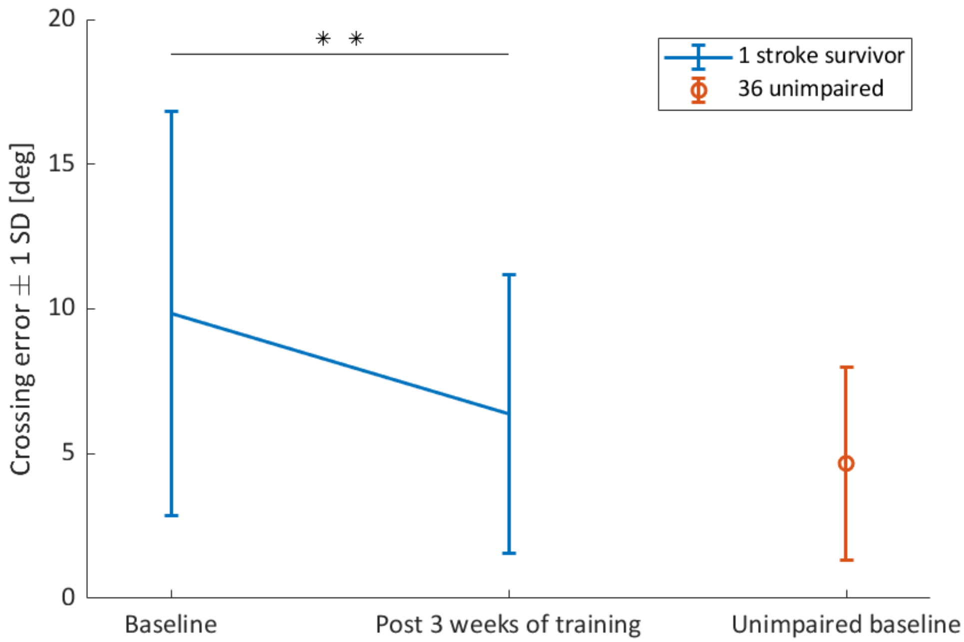

Figure 28. Improvement in proprioceptive acuity due to Propriopixels training and comparison to unimpaired baseline proprioceptive acuity.

The participant significantly improved (**$p < 0.01$) proprioceptive acuity after three weeks of Propriopixels training. Compared to the baseline Crisscross crossing errors presented in the previous chapter, at baseline the stroke survivor's proprioceptive acuity was 2.1 times greater (worse) than unimpaired participants, which decreased to 1.4 times greater after training.

Discussion

We evaluated the feasibility of P-Pong as a hand propriomotor rehabilitation activity after stroke. Specifically, we studied how to familiarize stroke survivors with a cognitively complex game that requires them to fuse multimodal sensory information, make cognitive gameplay decisions, and act on those decisions with corresponding finger movements and tested the game's training effects on upper extremity capacity and performance. The first participant played P-Pong in the least cognitively challenging mode (Video-only) and tested both physical and virtual assistance types, which yielded final success rate errors of 10.7%

and 3.8% respectively. Participant two played through a progression of Propriopixels familiarization conditions, beginning with three matches (60 trials) of assistance off and targets on, followed by seven matches of (140 trials) with targets and assistance on. Their crossing errors decreased within each condition, e.g. they began the no targets condition with a mean crossing error of 12.2 deg crossing error and ended with 5.6 deg. Turning targets on increased their mean crossing error to 7.4 deg, which decreased to 5.3 deg over the course of play. Participant three played through the same progression as two, only they began with a single Video-only display mode match (before playing Propriopixels). Their progression of crossing errors within each condition followed the same trend of initially increasing when targets were introduced, then decreasing with practice. Overall, both participants that played the Propriopixels display mode significantly decreased their crossing error from the first to last match ($p = 0.005$ and $p = 0.027$). The success control algorithm regulated success with two different assistance types (virtual where we varied paddle height and physical where we varied the level of forcing along an assistance trajectory), with mean final success rate errors of 10.7% (participant one, physical assist, 140 trials), 3.8% (participant one, virtual assist, 80 trials), 18.9% (participant two, physical assist, 140 trials), and 9.4% (participant three, physical assist, 180 trials). Turning the success control algorithm off and raising the difficulty back to its initial level increased success rate errors by 6.8% and 3.4% for participants two and three respectively, also indicating that the algorithm regulated success over a relatively small number of trials (80 – 180 across all conditions, compared to 320 in the previous chapter), indicating that both assistance types are capable of regulating success during game play. Lastly, participant one performed nine hours of Propriopixels training over three weeks. Their passive proprioceptive acuity significantly improved ($p < 0.01$), their Motor Activity Log Amount of use and movement quality improved by clinically meaningful amounts, and their Box and Blocks Test scores were relatively stationary. We next discuss these results and a set of derived recommendations for familiarizing stroke survivors with multimodal rehabilitation games.

A. Did the participants understand P-Pong?

The primary aim of this study was to evaluate the extent to which chronic stroke survivors could understand and play P-Pong game. To test this, we stepped through a progression of game settings with two of the participants that culminated with the most complex configuration of game settings: Propriopixels with targets. During our first Propriopixels familiarization attempt with participant two, we noticed that it took several matches to understand the game. – it was difficult to ascertain from observing them alone, and they were aphasic and spoke very little. The progression of their crossing errors and difficulty indicates that their understanding improved by the end of play (after 260 trials total). Comparing their first and last match which both have the same difficulty and assistance off, they decreased their mean crossing error significantly improvement. This is impressive considering that their last match was played with targets and the first match was not. Analyzing the progression of difficulty (automatically adjusted by the success control algorithm), however, may indicate that participant two had trouble comprehending Propriopixels. When we turned the success control algorithm on it maximized assistance quickly and assistance remained high (Figure 25). Increasing assistance decreased position errors, however when assistance was switched off for the final three matches there was only a small increase in crossing error. Given the short period over which these crossing error changes occurred, this may indicate that physical assistance was compensating for a lack of game comprehension as opposed to compensating for a lack of propriomotor capacity. To address this, we began participant three's familiarization progression with the Video-only display mode, and we observed key differences in their experience. Despite having less upper extremity capacity than participant two (Box and Blocks Test 34 versus 49, respectively), participant three did not reach maximum assistance while they played Propriopixels with targets – they reached a minimum difficulty of 0.5 compared to 0.0 for participant two. This suggests that beginning Propriopixels familiarization with Video-only play may have helped participant three learn Propriopixels more quickly than participant two. Even so, both participants significantly decreased crossing errors from the first to the last match indicating that both were able to play and understand the most challenging P-Pong configuration - Propriopixels display mode with targets.

B. Recommendations for designing rehabilitation games

1. Use a "challenging" range of motion. Participant one grew increasingly frustrated as they were continually unable to extend their index finger and hit the top target. As their frustration grew their tone increased, creating a vicious cycle of missing targets. Although they were much more successful and content after we decreased their extension range of motion, this could have limited challenge in the "wrong" ways and ultimately could have limited training efficacy. To balance these variables of demotivation and challenge, we recommend customizing range of motion to a safe but challenging level for each participant, which may be outside their reachable workspace.

2. Scaffold the familiarization procedure. We gradually increased P-Pong complexity by beginning with the simplest, most intuitive game mode and incrementally introducing new

modes with multiple practice repetitions at each. In pedagogy this process is broadly referred to as "scaffolding". Bruner first defined scaffolding as "the steps taken to reduce the degrees of freedom in carrying out some tasks so that the child can concentrate on the difficult skill she is in the process of acquiring" [140]. Although our sample size was small, we believe increased scaffolding improved participant three's experience and resulted in better understanding of our most complex game mode.

3. Interleave varying levels of challenge. For the Propriopixels mode there was a variation in crossing error across the three targets, and that variation decreased with practice - an indicator of learning. Coupled with success control, these two ways of varying difficulty could be viewed as having faster (targets) and slower (success control algorithm) time constants, as the targets change more rapidly than the algorithm modulated difficulty level. These together could benefit learning by adjusting the overall challenge to the learner's ability's (success control algorithm), then randomly exposing them to different movement tasks (randomized targets) which has been shown in motor learning that random presentations of a task produce better learning than blocked presentations [141].

4. Success control – don't set it and forget it. Success is mediated by several factors such as activity understanding, attention, and motivation. We implemented the success control algorithm with the intention of compensating for the range of sensorimotor capacity of stroke survivors, however in this study alone we deduced that it may have (over)compensated for additional undesirable factors, namely understanding the game. This was not clear from observing game play alone. We recommend considering and analyzing

how any automatic control algorithms and game parameters could limit a participant's benefit from training.

Conclusion

We developed P-Pong to investigate the potential benefits of the novel sensory gaming paradigm that we call Propriopixels. For the first time, we evaluated the feasibility of playing P-Pong to train hand function after stroke. We found that two different types of assistance, virtual and physical, both regulated player success. Importantly, we found that participants were able to play and understand the most cognitively challenging configuration of the game and recommend taking care to design familiarization procedures that gradually build from very simple to complex versions of training activities. Propriopixels training significantly improved passive proprioceptive acuity and caused clinically meaningful increases in daily arm use. Following these positive results, our goal is to investigate the potential benefits of our multimodal target based Propriopixels training in chronic stroke survivors.

CHAPTER 6: CONCLUSIONS AND MAJOR CONTRIBUTIONS

I began this dissertation by identifying three important gaps that limit the design and effectiveness of upper extremity rehabilitation technology after stroke. First, we lack understanding of how impairment reduction can lead to use increase. Second, despite the prevalence of proprioceptive deficits after stroke and the potential role of proprioception in motor learning, there are no methods for intensely and engagingly training hand propriomotor capacity. Third, there is an unmet need for compact rehabilitation robotic devices suitable for home use. Through the statistical modeling, engineering development, and human subjects research described in this dissertation we made advances in precision rehabilitation, rehabilitation gaming paradigms, and rehabilitation robot design. We next summarize the key accomplishments of this work and discuss future research directions.

A. Precision rehabilitation

In Chapter 1 of this dissertation, we introduced the different perspectives of stroke rehabilitation and the concept of activity limitations. The International Classification of Functioning, Disability and Health developed by The World Health Organization in 2001, defines activity limitations as difficulties an individual may have in executing a task or action [7]. Activity limitations can be addressed in multiple ways. Compensation refers to using the less affected upper extremity to complete a task. Differently, restitution refers to the restoration of the upper extremity to original levels of capacity [142], [143]. However restoring capacity does not necessarily translate to increased use [53], which tends to lag motor and functional capacity gains [42]. The Threshold Hypothesis helps explain this trend, stating mechanistically that lag in use occurs until upper extremity capacity reaches a threshold at which a virtuous self-training cycle begins [10]. Toward the goal of catalyzing such virtuous self-training cycles, in Chapter 2 we investigated whether we could identify baseline measures that predict increases in daily arm use after therapy. We found that two simple clinical measures, the Box and Blocks Test and the Motor Activity Log Amount of Use subscale predicted whether participants increased their amount of use by at least one level on the Motor Activity Log Amount of Use subscale. The Box and Blocks Test was the strongest predictor, where increases in Box and Blocks Test scores raised the likelihood of increasing amount of use after therapy. Interestingly, baseline Motor Activity Log was inversely proportional – decreases in baseline use raised the likelihood of increasing use after therapy. Through this model we identified a relationship that we call “untapped use potential”, that those with the highest model predicted probabilities of increasing use had mismatched, low use relative to their capacity. Further, the type of therapy – technology-based, or conventional - did not explain whether participants increased use, suggesting that stroke survivors with untapped use potential may benefit from any reasonable, intensive therapy.

A practical future direction of this work is to identify stroke survivors who have untapped use potential and thus to predict who has a high likelihood of increasing upper extremity use. A next step toward that goal is to obtain many (hundreds) more samples of the Motor Activity Log and the Box and Blocks Test to build a predictive model. Following, we recommend investigating the type(s) of therapy that generates increases in use for stroke survivors with untapped use potential. Our results indicated that stroke survivors may benefit from any reasonable, intensive therapy. Perhaps those with untapped use potential could benefit from goal-directed, automated feedback through a wearable device – an intervention that is scalable, inexpensive, and can be deployed in a home environment.

For those that did not increase amount of use, did they increase in other measures? In Chapter 2 we quantified rehabilitation benefit as a binary increase in Motor Activity Log Amount of Use. Toward the goal of building models to inform the development of precision rehabilitation, a separate line of research could be to formulate a dependent variable that encompasses multiple dimensions of improvement - a simple example in this case being improving capacity and performance. This could be formulated in many ways, but a simple extension of our modeling work in Chapter 2 is to formulate it as multiclass classification problem, where the model estimates a likelihood of belonging to categories that correspond to increasing upper extremity capacity alone, performance alone, or both. A continuous analog of the multiclass model is to treat each category as a continuous dimension, which may be better suited for continuous, objective measures.

C. Rehabilitation gaming paradigms

In Chapter 3 of this dissertation, we identified a need for finger proprioception training strategies. There is mounting evidence that proprioception is an essential input for learning [99]: it was a strong behavioral predictor of hand motor gains following constrain-induced therapy [100] and robotic hand therapy in chronic stroke [41], [71], suggesting that the training and improvement of proprioception could improve motor learning and recovery [101]. Considering the prevalence of proprioceptive deficits after stroke [33], [144], [145], the large number of practice repetitions required for sensory [13] and motor [11], [12] learning, the few number of repetitions practiced during therapy sessions [14], [14], and the general lack of a clinical arsenal for proprioceptively-focused rehabilitation training, there is a need for engaging, intense training strategies that target finger proprioception. We developed a novel gaming paradigm "Propriopixels" to target finger proprioception training and implemented it in an interpretation of the classic Pong arcade game that we call Proprioceptive-Pong (P-Pong). In Chapter 3 we showed that playing P-Pong for 15 minutes significantly improved passive finger proprioception acuity, despite performing less repetitions and being less successful than the control group who played the game in a classic video game form and did not improve. With the goal of enhancing its training benefits, in Chapter 4 we implemented two ways of modulating in-game difficulty. Namely, we created an updated version of a previously-developed success control algorithm which regulated success completely virtually (without physically assisting the player), and we added a second game component of ball targets which required the player to hit the ball with visually cued paddle zones that correspond to finger 'poses' (index finger above, matched, or below the middle finger). We demonstrated that we were able to regulate success completely virtually without the physical assistance capabilities of a complex robotic device. Interestingly we found that although players significantly reduced position errors during game play, those improvements did not transfer to Crisscross, a robotic measure of passive proprioception acuity, which may be attributed to a Crisscross assessment floor effect and/or a specificity of learning to movement type and highlights avenues for future work. Next in Chapter 5, considering the complexity of the P-Pong game, we evaluated the extent to which stroke survivors could understand and play P-Pong. We found that a scaffolded progression of game settings that gradually increased in complexity significantly reduced errors over a short single session of play, indicating that participants could both understand and play the most complex configuration of the P-Pong game. We find this impressive considering that P-Pong conveys information simultaneously through separate afferent pathways; i.e. vision and proprioception. Lastly, we evaluated the effects of intense, multi-session Propriopixels training on propriomotor capacity and arm use with a single stroke survivor. We found that Propriopixels training significantly improved passive proprioceptive acuity and caused clinically meaningful increases in daily arm use indicating that Propriopixels training can improve arm use after stroke.

The natural progression of this work is to next study the benefits of Propriopixels based P-Pong training in a randomized controlled trial comprised of several training sessions spanning multiple weeks, compared with standard video training and proprioception-targeted training (Propriopixels) groups. The potential Crisscross assessment floor effect identified in Chapter 4, where P-Pong training improvements did not transfer to passive proprioceptive acuity measured by Crisscross, could be addressed by implementing the elements of the PINKIE Crisscross assessment that were potentially more proprioceptively challenging: varying the position where the fingers cross, moving each finger at an independent speed, and increasing the fastest tested speed. The potential effects of specificity of training identified in Chapter 4 could be addressed in multiple ways depending on the goals of the study. Without evidence of the potentially independent roles of passive and active position sense in upper extremity capacity and performance, the 'kitchen sink' approach would be to train both by playing P-Pong in the control modes of Chapters 3 (passive movement) and 4 (active movement). Alternatively, to study the independent effects of each, passive and active training could be split into separate study groups. In both the former 'kitchen sink' training or the latter, we recommend including both passive and active measures of proprioception acuity to continue to study the potential specificity of movement type to proprioceptive learning.

B. Rehabilitation robot design

In Chapter 1 of this dissertation, we highlighted the need for robotic devices in stroke rehabilitation. Because a large number of repetitions are required for learning [11]–[13], and a small number of repetitions are practiced during therapy sessions [14], home-based therapy is a cornerstone of stroke rehabilitation. Many rehabilitation robots are bulky, expensive, and simply not designed for home use. And yet, only a fraction of subacute stroke survivors may be capable of performing rehabilitation exercises with only sensor-based device [23]. The Propriopixels gaming paradigm gave rise to a simple robot design pattern, a “binary impedance robot”, that we presented in Chapter 3. Instead of rendering a continuum of impedances that require costly, bulky actuators and sensors, a binary impedance robot renders the impedance limits with a low cost, stiff actuator and clutch. We developed the binary impedance robot PINKIE to play P-Pong. It is capable of both actuating and sensing a multi degree of freedom finger furling/unfurling movement with a simple prismatic-revolute mechanism and position dependent clutch. Of note, we manufactured PINKIE entirely from 3D printed, laser cut, and off-the-shelf hobby grade components; it weighs <5 lbs and it easily fits on a tabletop. It is a proof-of-concept for a robotic device that is simple and safe enough to administer sensitive, reliable assessments and deliver engaging gamified training in a package that could be feasibly deployed in a clinic or home. And, although we transitioned to using the more complex FINGER robot in Chapters 4 and 5 in this dissertation, all the P-Pong updates detailed in Chapter 4 (targets and virtual success control) are playable with a simple binary impedance robot like PINKIE. In future work, our continued investigation of the Propriopixels paradigm will inform how simple binary impedance robotic devices like PINKIE should be designed.

A simple extension of the binary impedance robot that may open new applications in stroke rehabilitation is series elasticity. Borrowing from the field of lower extremity assistive devices where series elastic actuators are more commonly used, an elastic element and force sensor would allow for added functionality like gradable, physical assistance. In fact, this could be easily implemented in PINKIE's clutch design. In its current form, the clutch connects through magnetic force: with a permanent magnet affixed to its front, the linear actuator simply approaches the finger mechanism and when it gets close enough (< 0.5 inches), the finger mechanism is attracted to the actuator. If the permanent magnet were replaced with an electromagnet, then the direction and magnitude of the coupling force could be controlled in real-time to either attract or repel the finger mechanism and modulate the force on the user's finger. In this way, the actuator would act as a "shuttle" to maintain a distance in which the electromagnetic can effectively attract and repel the finger mechanism.

REFERENCES

[1] V. L. Feigin *et al.*, "Global, regional, and national burden of stroke and its risk factors, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019," *Lancet Neurol.*, vol. 20, no. 10, pp. 795–820, Oct. 2021, doi: 10.1016/S1474-4422(21)00252-0.

[2] E. S. Donkor, "Stroke in the 21st Century: A Snapshot of the Burden, Epidemiology, and Quality of Life," *Stroke Res. Treat.*, vol. 2018, p. 3238165, Nov. 2018, doi: 10.1155/2018/3238165.

[3] J. W. Sturm, H. M. Dewey, G. A. Donnan, R. A. L. Macdonell, J. J. McNeil, and A. G. Thrift, "Handicap After Stroke: How Does It Relate to Disability, Perception of Recovery, and Stroke Subtype?: The North East Melbourne Stroke Incidence Study (NEMESIS)," *Stroke*, vol. 33, no. 3, pp. 762–768, Mar. 2002, doi: 10.1161/hs0302.103815.

[4] M. González-Fernández and D. Feldman, "Rehabilitation of the Stroke Patient," in *Conn's Current Therapy 2021*, R. D. Kellerman and D. Rakel, Eds. Elsevier Health Sciences, 2021.

[5] Consensus Panel on the Management of the Hemiplegic Arm and Hand, *Management of the post stroke arm and hand: Treatment recommendations of the 2001 consensus panel*. Heart and Stroke Foundation of Ontario, 2001.

[6] E. L. Miller *et al.*, "Comprehensive Overview of Nursing and Interdisciplinary Rehabilitation Care of the Stroke Patient: A Scientific Statement From the American Heart Association," *Stroke*, vol. 41, no. 10, pp. 2402–2448, Oct. 2010, doi: 10.1161/STR.0b013e3181e7512b.

[7] "International Classification of Functioning, Disability and Health (ICF)." https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health (accessed Jul. 21, 2022).

[8] P. Langhorne, J. Bernhardt, and G. Kwakkel, "Stroke rehabilitation," *The Lancet*, vol. 377, no. 9778, pp. 1693–1702, May 2011, doi: 10.1016/S0140-6736(11)60325-5.

[9] C. J. Winstein *et al.*, "Guidelines for Adult Stroke Rehabilitation and Recovery: A Guideline for Healthcare Professionals From the American Heart Association/American Stroke Association," *Stroke*, vol. 47, no. 6, Jun. 2016, doi: 10.1161/STR.0000000000000098.

[10] N. Schweighofer, C. E. Han, S. L. Wolf, M. A. Arbib, and C. J. Winstein, "A Functional Threshold for Long-Term Use of Hand and Arm Function Can Be Determined: Predictions From a Computational Model and Supporting Data From the Extremity Constraint- Induced Therapy Evaluation (EXCITE) Trial," p. 10, 2009.

[11] Kwakkel Gert, Wagenaar Robert C., Koelman Tim W., Lankhorst Gustaaf J., and Koetsier Johan C., "Effects of Intensity of Rehabilitation After Stroke," *Stroke*, vol. 28, no. 8, pp. 1550–1556, Aug. 1997, doi: 10.1161/01.STR.28.8.1550.

[12] A. Abdullahi, “Effects of Number of Repetitions and Number of Hours of Shaping Practice during Constraint-Induced Movement Therapy: A Randomized Controlled Trial,” *Neurol. Res. Int.*, vol. 2018, pp. 1–9, Apr. 2018, doi: 10.1155/2018/5496408.

[13] S. Doyle, S. Bennett, S. E. Fasoli, and K. T. McKenna, “Interventions for sensory impairment in the upper limb after stroke,” *Cochrane Database Syst. Rev.*, Jun. 2010, doi: 10.1002/14651858.CD006331.pub2.

[14] C. E. Lang, J. R. MacDonald, and C. Gnip, “COUNTING REPS: AN OBSERVATIONAL STUDY OF OUTPATIENT DAY TREATMENT FOR PEOPLE WITH HEMIPARESIS,” *J. Neurol. Phys. Ther.*, vol. 30, no. 4, p. 209, Dec. 2006, doi: 10.1097/01.NPT.0000281300.86409.c4.

[15] M. Shaughnessy, B. M. Resnick, and R. F. Macko, “Testing a Model of Post-Stroke Exercise Behavior,” *Rehabil. Nurs.*, vol. 31, no. 1, pp. 15–21, 2006, doi: https://doi.org/10.1002/j.2048-7940.2006.tb00005.x.

[16] S. Patel, H. Park, P. Bonato, L. Chan, and M. Rodgers, “A review of wearable sensors and systems with application in rehabilitation,” *J. NeuroEngineering Rehabil.*, vol. 9, no. 1, p. 21, Dec. 2012, doi: 10.1186/1743-0003-9-21.

[17] M. Noorkõiv, H. Rodgers, and C. I. Price, “Accelerometer measurement of upper extremity movement after stroke: a systematic review of clinical studies,” *J. NeuroEngineering Rehabil.*, vol. 11, no. 1, p. 144, 2014, doi: 10.1186/1743-0003-11-144.

[18] Q. Wang, P. Markopoulos, B. Yu, W. Chen, and A. Timmermans, “Interactive wearable systems for upper body rehabilitation: a systematic review,” *J. NeuroEngineering Rehabil.*, vol. 14, no. 1, p. 20, Mar. 2017, doi: 10.1186/s12984-017-0229-y.

[19] K. Myllymaa, R. Raisamo, J. Lylykangas, J. Heikkinen, and V. Surakka, “RehApp – A Wearable Haptic System for Rehabilitation and Sports Training,” in *Haptics: Perception, Devices, Mobility, and Communication*, vol. 7283, P. Isokoski and J. Springare, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 210–213. doi: 10.1007/978-3-642-31404-9_38.

[20] I. cheol Jeong and J. Finkelstein, “Computer-assisted upper extremity training using interactive biking exercise (iBikE) platform,” in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug. 2012, pp. 6095–6099. doi: 10.1109/EMBC.2012.6347384.

[21] A. Holden *et al.*, “CueS: cueing for upper limb rehabilitation in stroke,” in *Proceedings of the 2015 British HCI Conference*, Lincoln Lincolnshire United Kingdom, Jul. 2015, pp. 18–25. doi: 10.1145/2783446.2783576.

[22] M. A. Gull, S. Bai, and T. Bak, “A Review on Design of Upper Limb Exoskeletons,” *Robotics*, vol. 9, no. 1, Art. no. 1, Mar. 2020, doi: 10.3390/robotics9010016.

[23] Q. Sanders, V. Chan, R. Augsburger, S. C. Cramer, D. J. Reinkensmeyer, and A. H. Do, “Feasibility of Wearable Sensing for In-Home Finger Rehabilitation Early After Stroke,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 6, pp. 1363–1372, Jun. 2020, doi: 10.1109/TNSRE.2020.2988177.

[24] L. Proteau, "Chapter 4 On The Specificity of Learning and the Role of Visual Information for Movement Control," in *Advances in Psychology*, vol. 85, L. Proteau and D. Elliott, Eds. North-Holland, 1992, pp. 67–103. doi: 10.1016/S0166-4115(08)62011-7.

[25] L. Proteau and G. Isabelle, "On the Role of Visual Afferent Information for the Control of Aiming Movements Toward Targets of Different Sizes," *J. Mot. Behav.*, vol. 34, no. 4, pp. 367–384, Dec. 2002, doi: 10.1080/00222890209601954.

[26] C. Robin, L. Toussaint, Y. Blandin, and L. Proteau, "Specificity of Learning in a Video-Aiming Task: Modifying the Salience of Dynamic Visual Cues," *J. Mot. Behav.*, vol. 37, no. 5, pp. 367–376, Sep. 2005, doi: 10.3200/JMBR.37.5.367-376.

[27] L. Proteau, "Visual afferent information dominates other sources of afferent information during mixed practice of a video-aiming task," *Exp. Brain Res.*, vol. 161, no. 4, pp. 441–456, Mar. 2005, doi: 10.1007/s00221-004-2090-z.

[28] L. M. Carey and T. A. Matyas, "Training of Somatosensory Discrimination After Stroke: Facilitation of Stimulus Generalization," *Am. J. Phys. Med. Rehabil.*, vol. 84, no. 6, pp. 428–442, Jun. 2005, doi: 10.1097/01.PHM.0000159971.12096.7F.

[29] D. J. Reinkensmeyer *et al.*, "Computational neurorehabilitation: modeling plasticity and learning to predict recovery," *J. NeuroEngineering Rehabil.*, vol. 13, no. 1, p. 42, Dec. 2016, doi: 10.1186/s12984-016-0148-3.

[30] D. Rand, D. Gottlieb, and P. L. (Tamar) Weiss, "Recovery of Patients with a Combined Motor and Proprioception Deficit During the First Six Weeks of Post Stroke Rehabilitation," *Phys. Occup. Ther. Geriatr.*, vol. 18, no. 3, pp. 69–87, Jan. 2001, doi: 10.1080/J148v18n03_05.

[31] S. Hunter and P. Crome, "Hand function and stroke," *Rev. Clin. Gerontol.*, vol. 12, no. 1, pp. 68–81, Feb. 2002, doi: 10.1017/S0959259802012194.

[32] M. Yekutiel, *Sensory re-education of the hand after stroke / Margaret Yekutiel.* London ; Whurr, 2000.

[33] L. M. Carey, "Somatosensory Loss after Stroke," *Crit. Rev. Phys. Rehabil. Med.*, vol. 7, no. 1, 1995, doi: 10.1615/CritRevPhysRehabilMed.v7.i1.40.

[34] E. Moberg, "Two-point discrimination test. A valuable part of hand surgical rehabilitation, e.g. in tetraplegia," *Scand. J. Rehabil. Med.*, vol. 22, no. 3, pp. 127–134, Jan. 1990.

[35] D. T. Wade, R. Langton-Hewer, V. A. Wood, C. E. Skilbeck, and H. M. Ismail, "The hemiplegic arm after stroke: measurement and recovery.," *J. Neurol. Neurosurg. Psychiatry*, vol. 46, no. 6, pp. 521–524, Jun. 1983, doi: 10.1136/jnnp.46.6.521.

[36] J. E. Aman, N. Elangovan, I.-L. Yeh, and J. Konczak, "The effectiveness of proprioceptive training for improving motor function: a systematic review," *Front. Hum. Neurosci.*, vol. 8, Jan. 2015, doi: 10.3389/fnhum.2014.01075.

[37] C. T. Fuentes and A. J. Bastian, "Where Is Your Arm? Variations in Proprioception Across Space and Tasks," *J. Neurophysiol.*, vol. 103, no. 1, pp. 164–171, Jan. 2010, doi: 10.1152/jn.00494.2009.

[38] N. Elangovan, I.-L. Yeh, J. Holst-Wolf, and J. Konczak, “A robot-assisted sensorimotor training program can improve proprioception and motor function in stroke survivors,” in *2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR)*, Toronto, ON, Canada, Jun. 2019, pp. 660–664. doi: 10.1109/ICORR.2019.8779409.

[39] A. H. Kobeissi, G. Lanza, R. Berta, F. Bellotti, and A. D. Gloria, “Development of a Hardware/Software System for Proprioception Exergaming,” *Int. J. Serious Games*, vol. 5, no. 2, Art. no. 2, Jun. 2018, doi: 10.17083/ijsg.v5i2.244.

[40] A. I. R. Kottink, L. van Velsen, J. Wagenaar, and J. H. Buurke, “Assessing the gaming experience of a serious exergame for balance problems: Results of a preliminary study,” in *2015 International Conference on Virtual Rehabilitation (ICVR)*, Jun. 2015, pp. 135–136. doi: 10.1109/ICVR.2015.7358614.

[41] M. L. Ingemanson, J. R. Rowe, V. Chan, E. T. Wolbrecht, D. J. Reinkensmeyer, and S. C. Cramer, “Somatosensory system integrity explains differences in treatment response after stroke,” *Neurology*, vol. 92, no. 10, pp. e1098–e1108, Mar. 2019, doi: 10.1212/WNL.0000000000007041.

[42] D. Rand and J. J. Eng, “Predicting Daily Use of the Affected Upper Extremity 1 Year after Stroke,” *J. Stroke Cerebrovasc. Dis.*, vol. 24, no. 2, pp. 274–283, Feb. 2015, doi: 10.1016/j.jstrokecerebrovasdis.2014.07.039.

[43] E. Flores, G. Tobon, E. Cavallaro, F. I. Cavallaro, J. C. Perry, and T. Keller, “Improving patient motivation in game development for motor deficit rehabilitation,” in *Proceedings of the 2008 International Conference in Advances on Computer Entertainment Technology - ACE ’08*, Yokohama, Japan, 2008, p. 381. doi: 10.1145/1501750.1501839.

[44] D. K. Zondervan *et al.*, “Home-based hand rehabilitation after chronic stroke: Randomized, controlled single-blind trial comparing the MusicGlove with a conventional exercise program,” *J. Rehabil. Res. Dev.*, vol. 53, no. 4, pp. 457–472, 2016, doi: 10.1682/JRRD.2015.04.0057.

[45] J. H. van der Lee, I. A. Snels, H. Beckerman, G. J. Lankhorst, R. C. Wagenaar, and L. M. Bouter, “Exercise therapy for arm function in stroke patients: a systematic review of randomized controlled trials,” *Clin. Rehabil.*, vol. 15, no. 1, pp. 20–31, Feb. 2001, doi: 10.1191/026921501677557755.

[46] M. A. Guadagnoli and T. D. Lee, “Challenge Point: A Framework for Conceptualizing the Effects of Various Practice Conditions in Motor Learning,” *J. Mot. Behav.*, vol. 36, no. 2, pp. 212–224, Jul. 2004, doi: 10.3200/JMBR.36.2.212-224.

[47] G. Wulf and R. Lewthwaite, “Optimizing performance through intrinsic motivation and attention for learning: The OPTIMAL theory of motor learning,” *Psychon. Bull. Rev.*, vol. 23, no. 5, pp. 1382–1414, Oct. 2016, doi: 10.3758/s13423-015-0999-9.

[48] S. S. Virani *et al.*, “Heart Disease and Stroke Statistics—2020 Update: A Report From the American Heart Association,” *Circulation*, vol. 141, no. 9, Mar. 2020, doi: 10.1161/CIR.0000000000000757.

[49] A. D. Lopez, C. D. Mathers, M. Ezzati, D. T. Jamison, and C. J. L. Murray, “Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data,” *Lancet Lond. Engl.*, vol. 367, no. 9524, pp. 1747–1757, May 2006, doi: 10.1016/S0140-6736(06)68770-9.

[50] C. P. Warlow, “Epidemiology of stroke,” *The Lancet*, vol. 352, pp. S1–S4, Oct. 1998, doi: 10.1016/S0140-6736(98)90086-1.

[51] R. de Haan, N. Aaronson, M. Limburg, R. L. Hewer, and H. van Crevel, “Measuring quality of life in stroke.,” *Stroke*, vol. 24, no. 2, pp. 320–327, Feb. 1993, doi: 10.1161/01.STR.24.2.320.

[52] K. Kim, Y. M. Kim, and E. K. Kim, “Correlation between the Activities of Daily Living of Stroke Patients in a Community Setting and Their Quality of Life,” *J. Phys. Ther. Sci.*, vol. 26, no. 3, pp. 417–419, 2014, doi: 10.1589/jpts.26.417.

[53] P. Raghavan, “Upper Limb Motor Impairment After Stroke,” *Phys. Med. Rehabil. Clin. N. Am.*, vol. 26, no. 4, pp. 599–610, Nov. 2015, doi: 10.1016/j.pmr.2015.06.008.

[54] S.-Y. Chen and C. J. Winstein, “A Systematic Review of Voluntary Arm Recovery in Hemiparetic Stroke: Critical Predictors for Meaningful Outcomes Using the International Classification of Functioning, Disability, and Health,” *J. Neurol. Phys. Ther.*, vol. 33, no. 1, pp. 2–13, Mar. 2009, doi: 10.1097/NPT.0b013e318198a010.

[55] F. Coupar, A. Pollock, P. Rowe, C. Weir, and P. Langhorne, “Predictors of upper limb recovery after stroke: a systematic review and meta-analysis,” *Clin. Rehabil.*, vol. 26, no. 4, pp. 291–313, Apr. 2012, doi: 10.1177/0269215511420305.

[56] L. Kwah and R. Herbert, “Prediction of Walking and Arm Recovery after Stroke: A Critical Review,” *Brain Sci.*, vol. 6, no. 4, p. 53, Nov. 2016, doi: 10.3390/brainsci6040053.

[57] S. Prabhakaran *et al.*, “Inter-individual Variability in the Capacity for Motor Recovery After Ischemic Stroke,” *Neurorehabil. Neural Repair*, vol. 22, no. 1, pp. 64–71, Jan. 2008, doi: 10.1177/1545968307305302.

[58] E. Zarahn *et al.*, “Prediction of Motor Recovery Using Initial Impairment and fMRI 48 h Poststroke,” *Cereb. Cortex*, vol. 21, no. 12, pp. 2712–2721, Dec. 2011, doi: 10.1093/cercor/bhr047.

[59] C. Winters, E. E. H. van Wegen, A. Daffertshofer, and G. Kwakkel, “Generalizability of the Proportional Recovery Model for the Upper Extremity After an Ischemic Stroke,” *Neurorehabil. Neural Repair*, vol. 29, no. 7, pp. 614–622, Aug. 2015, doi: 10.1177/1545968314562115.

[60] R. L. Hawe, S. H. Scott, and S. P. Dukelow, “Taking Proportional Out of Stroke Recovery,” *Stroke*, vol. 50, no. 1, pp. 204–211, Jan. 2019, doi: 10.1161/STROKEAHA.118.023006.

[61] T. M. H. Hope, K. Friston, C. J. Price, A. P. Leff, P. Rotshtein, and H. Bowman, “Recovery after stroke: not so proportional after all?,” *Brain*, vol. 142, no. 1, pp. 15–22, Jan. 2019, doi: 10.1093/brain/awy302.

[62] M. R. Senesh and D. J. Reinkensmeyer, “Breaking Proportional Recovery After Stroke,” *Neurorehabil. Neural Repair*, p. 154596831986871, Aug. 2019, doi: 10.1177/1545968319868718.

[63] W. D. Byblow, C. M. Stinear, P. A. Barber, M. A. Petoe, and S. J. Ackerley, “Proportional recovery after stroke depends on corticomotor integrity,” *Ann. Neurol.*, vol. 78, no. 6, pp. 848–859, 2015, doi: 10.1002/ana.24472.

[64] S.-W. Park, S. L. Wolf, S. Blanton, C. Winstein, and D. S. Nichols-Larsen, “The EXCITE Trial: Predicting a Clinically Meaningful Motor Activity Log Outcome,” *Neurorehabil. Neural Repair*, vol. 22, no. 5, pp. 486–493, Sep. 2008, doi: 10.1177/1545968308316906.

[65] C. E. Lang, K. J. Waddell, J. Barth, C. L. Holleran, M. J. Strube, and M. D. Bland, “Upper Limb Performance in Daily Life Approaches Plateau Around Three to Six Weeks Post-stroke,” *Neurorehabil. Neural Repair*, vol. 35, no. 10, pp. 903–914, Oct. 2021, doi: 10.1177/15459683211041302.

[66] C. B. Lundquist, J. F. Nielsen, and I. C. Brunner, “Prediction of Upper Limb use Three Months after Stroke: A Prospective Longitudinal Study,” *J. Stroke Cerebrovasc. Dis.*, vol. 30, no. 11, p. 106025, Nov. 2021, doi: 10.1016/j.jstrokecerebrovasdis.2021.106025.

[67] Y. Li, W. Liao, Y. Hsieh, K. Lin, and C. Chen, “Predictors of Clinically Important Changes in Actual and Perceived Functional Arm Use of the Affected Upper Limb After Rehabilitative Therapy in Chronic Stroke,” *Arch. Phys. Med. Rehabil.*, vol. 101, no. 3, pp. 442–449, Mar. 2020, doi: 10.1016/j.apmr.2019.08.483.

[68] J. H. van der Lee, H. Beckerman, D. L. Knol, H. C. W. de Vet, and L. M. Bouter, “Clinimetric Properties of the Motor Activity Log for the Assessment of Arm Use in Hemiparetic Patients,” *Stroke*, vol. 35, no. 6, pp. 1410–1414, Jun. 2004, doi: 10.1161/01.STR.0000126900.24964.7e.

[69] S. Chen, S. L. Wolf, Q. Zhang, P. A. Thompson, and C. J. Winstein, “Minimal Detectable Change of the Actual Amount of Use Test and the Motor Activity Log: The EXCITE Trial,” *Neurorehabil. Neural Repair*, vol. 26, no. 5, pp. 507–514, Jun. 2012, doi: 10.1177/1545968311425048.

[70] M.-H. Milot *et al.*, “A crossover pilot study evaluating the functional outcomes of two different types of robotic movement training in chronic stroke survivors using the arm exoskeleton BONES,” *J. NeuroEngineering Rehabil.*, vol. 10, no. 1, p. 112, 2013, doi: 10.1186/1743-0003-10-112.

[71] J. B. Rowe, V. Chan, M. L. Ingemanson, S. C. Cramer, E. T. Wolbrecht, and D. J. Reinkensmeyer, “Robotic Assistance for Training Finger Movement Using a Hebbian Model: A Randomized Controlled Trial,” *Neurorehabil. Neural Repair*, vol. 31, no. 8, pp. 769–780, Aug. 2017, doi: 10.1177/1545968317721975.

[72] D. K. Zondervan, R. Augsburger, B. Bodenhoefer, N. Friedman, D. J. Reinkensmeyer, and S. C. Cramer, “Machine-Based, Self-guided Home Therapy for Individuals With Severe Arm Impairment After Stroke: A Randomized Controlled Trial,” *Neurorehabil. Neural Repair*, vol. 29, no. 5, pp. 395–406, Jun. 2015, doi: 10.1177/1545968314550368.

[73] B. W. Smith *et al.*, “Using a bimanual lever-driven wheelchair for arm movement practice early after stroke: A pilot, randomized, controlled, single-blind trial,” *Clin. Rehabil.*, vol. 35, no. 11, pp. 1577–1589, Nov. 2021, doi: 10.1177/02692155211014362.

[74] D. Schwerz de Lucena, “New Technologies for On-Demand Hand Rehabilitation in the Living Environment after Neurologic Injury,” UC Irvine, 2019. Accessed: Mar. 14, 2022. [Online]. Available: https://escholarship.org/uc/item/42423813

[75] A. M. Hammer and B. Lindmark, “Responsiveness and validity of the Motor Activity Log in patients during the subacute phase after stroke,” *Disabil. Rehabil.*, vol. 32, no. 14, pp. 1184–1193, Jan. 2010, doi: 10.3109/09638280903437253.

[76] G. Uswatte, E. Taub, D. Morris, K. Light, and P. A. Thompson, “The Motor Activity Log-28: Assessing daily use of the hemiparetic arm after stroke,” *Neurology*, vol. 67, no. 7, pp. 1189–1194, Oct. 2006, doi: 10.1212/01.wnl.0000238164.90657.c2.

[77] C. E. Lang, D. F. Edwards, R. L. Birkenmeier, and A. W. Dromerick, “Estimating Minimal Clinically Important Differences of Upper-Extremity Measures Early After Stroke,” *Arch. Phys. Med. Rehabil.*, vol. 89, no. 9, pp. 1693–1700, Sep. 2008, doi: 10.1016/j.apmr.2008.02.022.

[78] B. Essers, C. Biering Lundquist, G. Verheyden, and I. C. Brunner, “Determinants of Different Aspects of Upper-Limb Activity after Stroke,” *Sensors*, vol. 22, no. 6, p. 2273, Mar. 2022, doi: 10.3390/s22062273.

[79] S. L. Fritz, K. E. Light, S. N. Clifford, T. S. Patterson, A. L. Behrman, and S. B. Davis, “Descriptive Characteristics as Potential Predictors of Outcomes Following Constraint-Induced Movement Therapy for People After Stroke,” *Phys. Ther.*, vol. 86, no. 6, pp. 825–832, Jun. 2006, doi: 10.1093/ptj/86.6.825.

[80] H. Feys, W. D. Weerdt, G. Nuyens, A. V. D. Winckel, B. Selz, and C. Kiekens, “Predicting motor recovery of the upper limb after stroke rehabilitation: value of a clinical examination,” *Physiother. Res. Int.*, vol. 5, no. 1, pp. 1–18, 2000, doi: https://doi.org/10.1002/pri.180.

[81] M. Rijntjes *et al.*, “Individual Factors in Constraint-Induced Movement Therapy after Stroke,” *Neurorehabil. Neural Repair*, vol. 19, no. 3, pp. 238–249, Sep. 2005, doi: 10.1177/1545968305279205.

[82] V. Mathiowetz, G. Volland, N. Kashman, and K. Weber, “Adult Norills for the Box and Block Test of Manual Dexterity,” p. 6, Jun. 1985.

[83] G. Heinze, C. Wallisch, and D. Dunkler, “Variable selection – A review and recommendations for the practicing statistician,” *Biom. J.*, vol. 60, no. 3, pp. 431–449, 2018, doi: 10.1002/bimj.201700067.

[84] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning*, vol. 103. New York, NY: Springer New York, 2013. doi: 10.1007/978-1-4614-7138-7.

[85] F. E. Harrell , *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*. Cham: Springer International Publishing, 2015. doi: 10.1007/978-3-319-19425-7.

[86] R. M. Dannenbaum and L. A. Jones, "The Assessment and Treatment of Patients Who Have Sensory Loss Following Cortical Lesions," *J. Hand Ther.*, vol. 6, no. 2, pp. 130–138, Apr. 1993, doi: 10.1016/S0894-1130(12)80294-8.

[87] S. Hesse, C. Werner, M. Pohl, J. Mehrholz, U. Puzich, and H. I. Krebs, "Mechanical Arm Trainer for the Treatment of the Severely Affected Arm After a Stroke: A Single-Blinded Randomized Trial in Two Centers," *Am. J. Phys. Med. Rehabil.*, vol. 87, no. 10, pp. 779–788, Oct. 2008, doi: 10.1097/PHM.0b013e318186b4bc.

[88] S. Masiero, M. Armani, and G. Rosati, "Upper-limb robot-assisted therapy in rehabilitation of acute stroke patients: Focused review and results of new randomized controlled trial," *J. Rehabil. Res. Dev.*, vol. 48, no. 4, p. 355, 2011, doi: 10.1682/JRRD.2010.04.0063.

[89] H. C. Fischer, K. Stubblefield, T. Kline, X. Luo, R. V. Kenyon, and D. G. Kamper, "Hand Rehabilitation Following Stroke: A Pilot Study of Assisted Finger Extension Training in a Virtual Environment," *Top. Stroke Rehabil.*, vol. 14, no. 1, pp. 1–12, Jan. 2007, doi: 10.1310/tsr1401-1.

[90] J. Barth, J. W. Klaesner, and C. E. Lang, "Relationships between accelerometry and general compensatory movements of the upper limb after stroke," *J. NeuroEngineering Rehabil.*, vol. 17, no. 1, p. 138, Oct. 2020, doi: 10.1186/s12984-020-00773-4.

[91] K. S. Hayward, J. J. Eng, L. A. Boyd, B. Lakhani, J. Bernhardt, and C. E. Lang, "Exploring the Role of Accelerometers in the Measurement of Real World Upper-Limb Use After Stroke," *Brain Impair.*, vol. 17, no. 1, pp. 16–33, Mar. 2016, doi: 10.1017/BrImp.2015.21.

[92] D. Schwerz de Lucena, J. Rowe, V. Chan, and D. J. Reinkensmeyer, "Magnetically Counting Hand Movements: Validation of a Calibration-Free Algorithm and Application to Testing the Threshold Hypothesis of Real-World Hand Use after Stroke," *Sensors*, vol. 21, no. 4, Art. no. 4, Jan. 2021, doi: 10.3390/s21041502.

[93] J. A. Franck, R. J. E. M. Smeets, and H. A. M. Seelen, "Changes in actual arm-hand use in stroke patients during and after clinical rehabilitation involving a well-defined arm-hand rehabilitation program: A prospective cohort study," *PLOS ONE*, vol. 14, no. 4, p. e0214651, Apr. 2019, doi: 10.1371/journal.pone.0214651.

[94] R. J. Prescott, W. M. Garraway, and A. J. Akhtar, "Predicting functional outcome following acute stroke using a standard clinical examination.," *Stroke*, vol. 13, no. 5, pp. 641–647, Sep. 1982, doi: 10.1161/01.STR.13.5.641.

[95] E. Ekstrand, L. Rylander, J. Lexell, and C. Brogårdh, "Perceived ability to perform daily hand activities after stroke and associated factors: a cross-sectional study," *BMC Neurol.*, vol. 16, no. 1, p. 208, Nov. 2016, doi: 10.1186/s12883-016-0733-x.

[96] E. S. Lawrence *et al.*, "Estimates of the Prevalence of Acute Stroke Impairments and Disability in a Multiethnic Population," *Stroke*, vol. 32, no. 6, pp. 1279–1284, Jun. 2001, doi: 10.1161/01.STR.32.6.1279.

[97] S. S. Virani *et al.*, “Heart Disease and Stroke Statistics—2021 Update: A Report From the American Heart Association,” *Circulation*, vol. 143, no. 8, Feb. 2021, doi: 10.1161/CIR.0000000000000950.

[98] U. Sveen, E. Bautz-Holter, K. Margrethe Sodring, T. Bruun Wyller, and K. Laake, “Association between impairments, self-care ability and social activities 1 year after stroke,” *Disabil. Rehabil.*, vol. 21, no. 8, pp. 372–377, Jan. 1999, doi: 10.1080/096382899297477.

[99] J. D. Wong, D. A. Kistemaker, A. Chin, and P. L. Gribble, “Can proprioceptive training improve motor learning?,” *J. Neurophysiol.*, vol. 108, no. 12, pp. 3313–3321, Dec. 2012, doi: 10.1152/jn.00122.2012.

[100] S. L. Wolf *et al.*, “Effect of Constraint-Induced Movement Therapy on Upper Extremity Function 3 to 9 Months After StrokeThe EXCITE Randomized Clinical Trial,” *JAMA*, vol. 296, no. 17, pp. 2095–2104, Nov. 2006, doi: 10.1001/jama.296.17.2095.

[101] L. M. Carey, T. A. Matyas, and L. E. Oke, “Sensory loss in stroke patients: Effective training of tactile and proprioceptive discrimination,” *Arch. Phys. Med. Rehabil.*, vol. 74, no. 6, pp. 602–611, Jun. 1993, doi: 10.1016/0003-9993(93)90158-7.

[102] N. Lincoln, J. Crow, J. Jackson, G. Waters, S. Adams, and P. Hodgson, “The unreliability of sensory assessments,” *Clin. Rehabil.*, vol. 5, no. 4, pp. 273–282, Nov. 1991, doi: 10.1177/026921559100500403.

[103] M. Balaam *et al.*, “Motivating mobility: designing for lived motivation in stroke rehabilitation,” in *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI ’11*, Vancouver, BC, Canada, 2011, p. 3073. doi: 10.1145/1978942.1979397.

[104] N. Maclean and P. Pound, “A critical review of the concept of patient motivation in the literature on physical rehabilitation,” *Soc. Sci. Med. 1982*, vol. 50, pp. 495–506, Mar. 2000, doi: 10.1016/S0277-9536(99)00334-2.

[105] H. Taheri *et al.*, “Design and preliminary evaluation of the FINGER rehabilitation robot: controlling challenge and quantifying finger individuation during musical computer game play,” *J. NeuroEngineering Rehabil.*, vol. 11, no. 1, p. 10, 2014, doi: 10.1186/1743-0003-11-10.

[106] M. L. Ingemanson, J. B. Rowe, V. Chan, E. T. Wolbrecht, S. C. Cramer, and D. J. Reinkensmeyer, “Use of a robotic device to measure age-related decline in finger proprioception,” *Exp. Brain Res.*, vol. 234, no. 1, pp. 83–93, Jan. 2016, doi: 10.1007/s00221-015-4440-4.

[107] J. Han, G. Waddington, R. Adams, J. Anson, and Y. Liu, “Assessing proprioception: A critical review of methods,” *J. Sport Health Sci.*, vol. 5, no. 1, pp. 80–90, Mar. 2016, doi: 10.1016/j.jshs.2014.10.004.

[108] R. Bauer and A. Gharabaghi, “Reinforcement learning for adaptive threshold control of restorative brain-computer interfaces: a Bayesian simulation,” *Front. Neurosci.*, vol. 9, 2015, doi: 10.3389/fnins.2015.00036.

[109] H. L. O’Brien, P. Cairns, and M. Hall, “A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form,” *Int. J. Hum.-Comput. Stud.*, vol. 112, pp. 28–39, Apr. 2018, doi: 10.1016/j.ijhcs.2018.01.004.

[110] “Intrinsic Motivation Inventory (IMI) – selfdeterminationtheory.org.” https://selfdeterminationtheory.org/intrinsic-motivation-inventory/ (accessed Apr. 16, 2021).

[111] S. Dechaumont-Palacin *et al.*, “Neural Correlates of Proprioceptive Integration in the Contralesional Hemisphere of Very Impaired Patients Shortly After a Subcortical Stroke: An fMRI Study,” *Neurorehabil. Neural Repair*, vol. 22, no. 2, pp. 154–165, Mar. 2008, doi: 10.1177/1545968307307118.

[112] JohnP. Wann and SamF. Ibrahim, “Does limb proprioception drift?,” *Exp. Brain Res.*, vol. 91, no. 1, Oct. 1992, doi: 10.1007/BF00230024.

[113] R. C. Wilson, A. Shenhav, M. Straccia, and J. D. Cohen, “The Eighty Five Percent Rule for optimal learning,” *Nat. Commun.*, vol. 10, no. 1, Art. no. 1, Nov. 2019, doi: 10.1038/s41467-019-12552-4.

[114] D. S. Reinsdorf, E. E. Mahan, and D. J. Reinkensmeyer, “Proprioceptive Gaming: Making Finger Sensation Training Intense and Engaging with the P-Pong Game and PINKIE Robot,” in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, Nov. 2021, pp. 6715–6720. doi: 10.1109/EMBC46164.2021.9631041.

[115] F. R. Sarlegna and R. L. Sainburg, “The Roles of Vision and Proprioception in the Planning of Reaching Movements,” in *Progress in Motor Control*, vol. 629, D. Sternad, Ed. Boston, MA: Springer US, 2009, pp. 317–335. doi: 10.1007/978-0-387-77064-2_16.

[116] C. Ghez, M. Favilla, M. F. Ghilardi, J. Gordon, R. Bermejo, and S. Pullman, “Discrete and continuous planning of hand movements and isometric force trajectories:,” *Exp. Brain Res.*, vol. 115, no. 2, pp. 217–233, Jun. 1997, doi: 10.1007/PL00005692.

[117] C. Ghez, J. Gordon, and M. F. Ghilardi, “Impairments of reaching movements in patients without proprioception. II. Effects of visual information on accuracy,” *J. Neurophysiol.*, vol. 73, no. 1, pp. 361–372, Jan. 1995, doi: 10.1152/jn.1995.73.1.361.

[118] F. R. Sarlegna, G. M. Gauthier, C. Bourdin, J.-L. Vercher, and J. Blouin, “Internally driven control of reaching movements: A study on a proprioceptively deafferented subject,” *Brain Res. Bull.*, vol. 69, no. 4, pp. 404–415, Apr. 2006, doi: 10.1016/j.brainresbull.2006.02.005.

[119] J. D. Cooke, S. Brown, R. Forget, and Y. Lamarre, “Initial agonist burst duration changes with movement amplitude in a deafferented patient,” *Exp. Brain Res.*, vol. 60, no. 1, Sep. 1985, doi: 10.1007/BF00237030.

[120] M. Darainy, S. Vahdat, and D. J. Ostry, “Perceptual learning in sensorimotor adaptation,” *J. Neurophysiol.*, vol. 110, no. 9, pp. 2152–2162, Nov. 2013, doi: 10.1152/jn.00439.2013.

[121] L. Winter, Q. Huang, J. V. L. Sertic, and J. Konczak, “The Effectiveness of Proprioceptive Training for Improving Motor Performance and Motor Dysfunction: A Systematic Review,” *Front. Rehabil. Sci.*, vol. 3, p. 830166, Apr. 2022, doi: 10.3389/fresc.2022.830166.

[122] E. D. J. Ramos Muñoz *et al.*, “Using Large-Scale Sensor Data to Test Factors Predictive of Perseverance in Home Movement Rehabilitation: Optimal Challenge and Steady Engagement,” *Front. Neurol.*, vol. 13, p. 896298, Jun. 2022, doi: 10.3389/fneur.2022.896298.

[123] A. F. Baranes, P.-Y. Oudeyer, and J. Gottlieb, “The effects of task difficulty, novelty and the size of the search space on intrinsically motivated exploration,” *Front. Neurosci.*, vol. 8, 2014, Accessed: Jul. 31, 2022. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnins.2014.00317

[124] S. J. Spencer, “Movement Training and Post-Stroke Rehabilitation Using a Six Degree of Freedom Upper-Extremity Robotic Orthosis and Virtual Environment,” Ph.D., University of California, Irvine, United States -- California. Accessed: Jan. 13, 2022. [Online]. Available: https://www.proquest.com/docview/911792099/abstract

[125] J. Schell, *The art of game design: a book of lenses*, Reprinted. Amsterdam: Elsevier/Morgan Kaufmann, 2010.

[126] S. Rogers, *Level Up! The Guide to Great Video Game Design*, 1st ed. Wiley, 2010.

[127] K. Salen, K. S. Tekinbaş, and E. Zimmerman, *Rules of Play: Game Design Fundamentals*. MIT Press, 2004.

[128] A. Becker and D. Görlich, “What is Game Balancing? - An Examination of Concepts,” *ParadigmPlus*, vol. 1, no. 1, Art. no. 1, Apr. 2020, doi: 10.55969/paradigmplus.v1n1a2.

[129] L. K. Bush, U. Hess, G. Wolford, L. K. Bush, and G. Wolford, “Transformations for withinsubject designs: A Monte Carlo investigation,” *Psychol. Bull.*, pp. 79–88, 1993.

[130] D. Yekutieli and Y. Benjamini, “Resampling-based false discovery rate controlling multiple test procedures for correlated test statistics,” *J. Stat. Plan. Inference*, vol. 82, no. 1–2, pp. 171–196, Dec. 1999, doi: 10.1016/S0378-3758(99)00041-5.

[131] J.-C. Metzger, O. Lambercy, and R. Gassert, “High-fidelity rendering of virtual objects with the ReHapticKnob - novel avenues in robot-assisted rehabilitation of hand function,” in *2012 IEEE Haptics Symposium (HAPTICS)*, Mar. 2012, pp. 51–56. doi: 10.1109/HAPTIC.2012.6183769.

[132] O. Lambercy, L. Dovat, R. Gassert, E. Burdet, C. L. Teo, and T. Milner, “A Haptic Knob for Rehabilitation of Hand Function,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 15, no. 3, pp. 356–366, Sep. 2007, doi: 10.1109/TNSRE.2007.903913.

[133] J. L. Emken, R. Benitez, and D. J. Reinkensmeyer, “Human-robot cooperative movement training: Learning a novel sensory motor transformation during walking with robotic assistance-as-needed,” *J. NeuroEngineering Rehabil.*, vol. 4, no. 1, p. 8, Dec. 2007, doi: 10.1186/1743-0003-4-8.

[134] A. Duschau-Wicke, J. von Zitzewitz, A. Caprez, L. Lunenburger, and R. Riener, “Path Control: A Method for Patient-Cooperative Robot-Aided Gait Rehabilitation,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 18, no. 1, pp. 38–48, Feb. 2010, doi: 10.1109/TNSRE.2009.2033061.

[135] D. J. Reinkensmeyer, J. L. Emken, and S. C. Cramer, “Robotics, Motor Learning, and Neurologic Recovery,” *Annu. Rev. Biomed. Eng.*, vol. 6, no. 1, pp. 497–525, Aug. 2004, doi: 10.1146/annurev.bioeng.6.040803.140223.

[136] T. Nef, M. Mihelj, and R. Riener, “ARMin: a robot for patient-cooperative arm therapy,” *Med. Biol. Eng. Comput.*, vol. 45, no. 9, pp. 887–900, Sep. 2007, doi: 10.1007/s11517-007-0226-6.

[137] R. A. Schmidt and C. B. Walter, “Knowledge of Results and Motor Learning: A Review and Critical Reappraisal,” p. 32.

[138] L. Marchal-Crespo, S. McHughen, S. C. Cramer, and D. J. Reinkensmeyer, “The effect of haptic guidance, aging, and initial skill level on motor learning of a steering task,” *Exp. Brain Res.*, vol. 201, no. 2, pp. 209–220, Mar. 2010, doi: 10.1007/s00221-009-2026-8.

[139] R. Sigrist, G. Rauter, R. Riener, and P. Wolf, “Augmented visual, auditory, haptic, and multimodal feedback in motor learning: A review,” *Psychon. Bull. Rev.*, vol. 20, no. 1, pp. 21–53, Feb. 2013, doi: 10.3758/s13423-012-0333-8.

[140] J. Bruner, “The role of dialogue in language acquisition,” in *The Child’s Conception of Language*, vol. 2, A. Sinclair, R. Jarvella, and W. Levelt, Eds. Springer New York, 1978, pp. 241–256.

[141] J. Shea and R. Morgan, “Contextual interference effects on the acquisition, retention, and transfer of a motor skill,” *J. Exp. Psychol. [Hum. Learn.]*, vol. 5, p. 179, Mar. 1979, doi: 10.1037/0278-7393.5.2.179.

[142] L. J. Rothi and J. Horner, “Restitution and substitution: two theories of recovery with application to neurobehavioral treatment,” *J. Clin. Neuropsychol.*, vol. 5, no. 1, pp. 73–81, Feb. 1983, doi: 10.1080/01688638308401152.

[143] M. F. Levin, J. A. Kleim, and S. L. Wolf, “What Do Motor ‘Recovery’ and ‘Compensation’ Mean in Patients Following Stroke?,” *Neurorehabil. Neural Repair*, vol. 23, no. 4, pp. 313–319, May 2009, doi: 10.1177/1545968308328727.

[144] S. P. Dukelow *et al.*, “Quantitative Assessment of Limb Position Sense Following Stroke,” *Neurorehabil. Neural Repair*, vol. 24, no. 2, pp. 178–187, Feb. 2010, doi: 10.1177/1545968309345267.

[145] L. Connell, N. Lincoln, and K. Radford, “Somatosensory impairment after stroke: frequency of different deficits and their recovery,” *Clin. Rehabil.*, vol. 22, no. 8, pp. 758–767, Aug. 2008, doi: 10.1177/0269215508090674.