THE USE OF PHYSIOLOGICAL SIGNALS AND MOTOR PERFORMANCE METRICS IN TASK DIFFICULTY ADAPTATION: IMPROVING ENGAGEMENT IN ROBOT-ASSISTED MOVEMENT THERAPY

by

Navid Shirzad

B.Sc., University of Tehran, 2009

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

The Faculty of Graduate Studies

(Mechanical Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

April 2013

Abstract

Before robot-assisted therapy regimens can be included in clinical practice, one of the major challenges to overcome is maintaining the patient's engagement in the therapy during the lengthy functional recovery period. Game designers and psychologists have theorized the mechanics of sustaining an individual's engagement in a task. In a motor learning context, to maintain motivation to continue an exercise, one must be kept exercising at one's desirable difficulty by manipulation of the task challenge over the course of treatment. Thus, this work was aimed to design a robotic therapy regimen that can automatically adjust the difficulty to motivate users to continue with the exercise. The main contributions of this thesis are to develop a method to predict the user's desirable difficulty and validate the effects of adaptively adjusting a robotic exercise on the user's perception of the task.

The theory of desirable difficulty relies on three main factors: meaningful levels of difficulties, knowledge of the user's challenge preference, and positive effects of exercising a task under the desirable difficulty conditions. Studies to develop implementations of the first two factors in the context of an upper-limb reaching task were conducted, and investigated the effects of practicing this task under the desirable difficulty conditions.

The first study implemented five error amplification (EA) methods for a reaching task and validated that users perceive each with a different challenge level. In the second study, users' physiological and motor performance metrics were collected, as well as self-reports of the user's challenge preference after exercising with each of the EAs. The efficiency of different machine learning methods in predicting a user's challenge preference based on different combinations of physiological and motor performance attributes were analyzed. In the third study, the control group received EAs in predefined random order while the experimental group received EAs based on the predictions of the trained machine learning algorithm. The experimental group reported statistically significant higher scores on the metrics that assessed satisfaction, attentiveness, and willingness to continue the task. These results support the approach of designing a robotic system capable of adjusting exercises to prolong individuals' engagement in stroke therapy.

Preface

Studies and experiments described in this thesis were performed with the approval of the Clinical Research Ethics Board (CREB) at the University of British Columbia, under ethics application **#H11-01215** "Distortion-based Robot Interface".

The material presented in Chapter 1 of this thesis is available as two works: a non peer-reviewed position paper and an extended abstract. The author was responsible for performing literature review, developing software, conducting pilot studies and data analysis, and writing the manuscripts. The position paper was presented by the co-author at IEEE Symposium on Robot and Human Interactive Communication Ro-Man 2011. The extended abstract was accepted to the Biennial Meeting of the Canadian Society of Biomechanics/Societe Canadienne de Biomechanique (CSB/SBC) held in June 2012. Both of these works were published online and the authors hold the copy right.

H. F. M. Van der Loos, and N. Shirzad, "Therapy Robotics: The Power of the Interface to Motivate," Ro-Man 2011 – The 20th IEEE International Symposium on Robot and Human Interactive Communication, 2011.

N. Shirzad, and H. F.M. Van der Loos," Estimation of Desirable Difficulty in a Robotic Therapy Regimen Based on Performance and Affect," CSB-SCB 2012 – The 17th Biennial Meeting of the Canadian Society of Biomechanics/Societe Canadienne de Biomechanique, 2012.

Work presented in Chapter 2 is available as a conference paper. The author was responsible for performing literature review, developing software, conducting studies and data analysis, and writing the manuscript.

N. Shirzad, and H.F.M. Van der Loos, "Error amplification to promote motor learning and motivation in therapy robotics," In Proceedings of IEEE International Conference on Engineering in Medicine and Biology Society EMBC, 2012, pp. 3907-3910.

The results presented in Chapter 3, in part, are accepted as a full paper to IEEE International Conference on Rehabilitation Robotics ICORR 2013 titled "Adaptation of task difficulty in rehabilitation exercises based on the user's motor Performance and physiological responses". The author was responsible for performing literature review, developing software, conducting studies and data analysis, and writing the manuscript.

Aside from the works listed above, the author of this thesis was also involved in submission of a journal paper as the second author. The paper is titled “Video games and rehabilitation: using design principles to enhance patient engagement” and is currently under review (accepted pending minor reviews). This review paper was motivated directly by the work presented in this thesis.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I am indebted to many individuals for their support, help, and friendship along my journey to receiving this degree.

First and foremost, I would like to express my gratitude to Dr. Mike Van der Loos, my thesis supervisor, for his confidence-building trust in me and the time and effort he has committed to my education. I would also like to thank Dr. Elizabeth Croft, who kindly sat in all of my research meetings and provided invaluable knowledge and feedback. The diverse backgrounds of my two mentors made for an enjoyable learning experience.

I acknowledge the support I received from the UBC Mechanical Engineering Department (UBC MECH), National Science and Engineering Research Council of Canada (NSERC), Peter Wall Institute for Advanced Studies (PWIAS), and Institute for Computing, Information and Cognitive Systems (ICICIS). I wish to express my appreciation to the individuals who contributed their time and interest into this research by participating in my studies.

I was given the opportunity to supervise two undergraduate students in the CARIS lab. Sara Sheikholeslami and Arnold Yeung worked with me on the development and integration of the physiological signal processing code that was used in the second and third studies of this thesis. The CARIS lab members, past and present, are stimulating, challenging, and supportive colleagues and this thesis has benefited from my discussions with them. I appreciate the technical assistance I received from Dr. Willem Atsma in preparing the robotic manipulandum that was used in the studies presented in this thesis. Tina Hung and Bulmaro Valdez helped in a portion of the second study in collecting physiological data.

I should thank my friends: Elham, Ergun, Eric, AJung, Jenny, Ida, Behrouz, Jeff, and Graham. Vancouver feels like home when you are around. I would like to thank Payam Mirshams and Masoud Jalali. Although we are separated by vast lands and oceans, you constantly prove to me that true friendship goes beyond all borders. Cheers to y'all!

It is a hard job to raise four kids, turning them from tiny troublemakers of young age to successful adults. I am grateful to my kind parents, Nariman and Mahbobe, for all their love. As I grew up, I was blessed to have three older siblings. I would like to thank my sisters and brother, Soroor, Setare, and Saeed, who have always been supportive of me, family's little "tah-taghari".

This thesis is dedicated to my mother, Mahbobe, who had her first stroke when I was a freshman at college. This bit of family history and the fact that I knew how my mom felt about doing her therapy sessions was the main reason for me to start working on this thesis.

We are taking a step at a time, mom!

1 Introduction

1.1 Stroke: Prevalence and Impacts

A cerebrovascular accident, or stroke, is a localized reduction in blood supply to a region of the brain and causes neurons in the affected area to die. Thus, a stroke typically causes loss of brain function. Ischemic strokes, which account for 80% of the total incidence, are caused by a blockage of vessels from a blood clot and interruption of blood flow. About 20% of strokes are caused by the rupture of blood vessels in the brain (hemorrhagic strokes) [1]. The effects of a stroke mainly depend on the location and size of the blood flow disturbance. Strokes predominantly involve one of the cerebral hemispheres, a condition that can affect one's ability to perform basic functions such as to move, see, remember, speak, reason, read and/or write.

In North America, approximately 610,000 people experience their first stroke each year [2]. In Canada, over 50,000 strokes are reported each year and 300,000 Canadians are living with the effects of stroke [1]. Hospital and physician services, lost wages, and decreased productivity associated with stroke cost the Canadian economy $3.6 billion per year. The health care cost for stroke survivors was estimated $18.8 billion in the United States in 2008, while an additional cost of $15.5 billion is associated with the lost productivity and premature mortality caused by stroke [3].

According to the World Health Organization, stroke is the leading cause of disability worldwide, with 80% of first time strokes leading to upper-extremity impairment [4]. While a third of this population recovers with minor disability, the majority of stroke survivors have moderate to severe impairments. Since the majority of daily life activities involve use of one's upper limbs [5], effective rehabilitation for both gross and fine motor movements is crucially needed in order to return stroke survivors back to their independent daily lives.

1.2 Upper Extremity Motor Rehabilitation After Stroke and Use of Robots

About half of upper limb functional recovery after stroke is spontaneous. Any additional recovery can be achieved only from intensive, repetitive therapy over months of time, and by stimulating neuroplastic changes in the brain's motor control pathways. Physical therapy is the traditional method of treatment after the first stage of recovery, aiming to improve function in impaired individuals.

In physical therapy, a form of exercise treatment, a caregiver works one-on-one with a stroke survivor (Figure 1.1) with a focus on maximizing movement potential by restoring motor skills and passive range of motion. Physiotherapy takes advantage of the brain's neuroplasticity, the brain's ability to reorganize and rewire neural pathways, with techniques such as massed practice of skills and constraint-induced therapy. In massed practice therapy the patient executes thousands of repetitions of simple tasks to stimulate neuroplasticity in the brain and initiate the brain's reorganization of motor pathways. However, in constraint-induced therapy the patient's unaffected limb is restrained to prevent compensatory actions and to overcome the learned non-use of the affected limb. Increased time in therapy, and thus the number of repetitions of therapeutic tasks, is correlated with functional recovery [6, 7].

Figure 1.1: In conventional physiotherapy a care giver works one-on-one with a client to help maintain function. Photo credit Jupiterimages/Creatas/Getty Images.

The risk of having a stroke doubles every 10 years for people older than 55 [1]. The ageing trend of the population in developed countries, in addition to increasing the risk factors for stroke, also projects a reduction in the ratio of working-age people to elders from 6:1 today to 3:1 in the next 40 years [8]. To meet the growing need for effective physical therapy regimens, the use of robotics to develop unconventional therapies has gained popularity in the past three decades. As an example, a 12-week, robot-assisted, upper-limb stroke therapy program was shown to have similar cost and efficacy to a conventional program [9].

Most of the research in the field of robot-assisted rehabilitation has been focused on replicating conventional therapy paradigms (e.g., repetitive movement training or massed practice) with

robots [10]. The majority of applications in upper-extremity rehabilitation are dedicated to recovery of gross motor movement, mainly defined as reaching tasks using the shoulder and elbow. A well-known example is MIT-Manus [11] (commercially available as InMotion2; Interactive Motion Technologies, Inc., Boston, MA), which uses an active assistive strategy. This system provides low-impedance force cues to correct the motion of the patient's arm, in planar, straight-line, point-to-point reaching tasks.

Recovery of fine motor movement (e.g., finger movement and grasping) also plays a significant role in restoring daily life quality and independence of stroke survivors. Interactive Motion Technologies introduced InMotion3 as a wrist rehabilitation robot, which uses the same neurorehabilitation principles as InMotion2. Studies have demonstrated improvement of wrist and finger function using the upper-limb subcomponent of the Fugl-Meyer scale [12].

Alongside more conventional rehabilitation paradigms, such as active assistive/resistive strategies, bilateral training of the paretic and non-paretic arms is another strategy. The Mirror Image Movement Enabler (MIME) system was among the first robotic systems that used this approach, guiding the paretic arm to move along the same 3D trajectory, in mirror image to the non-paretic arm. A group of subjects training with this system showed larger improvement in the proximal movement portion of the Fugl-Meyer test compared with a control group undergoing conventional therapy [13]. The MIME system has recently been augmented with the hand-function-exercise, RiceWrist [14], and is undergoing clinical trials similar to those using the inMotion3 system.

One of the advantages of robotic systems over a human therapist is that robots can be programmed to deliver unconventional therapies. These include use of virtual environments and gaming interfaces, distortion of visual feedback and augmentation of error via force fields. The GENTLE/s system uses a HapticMASTER robot (MOOG FCS Robotics Inc., Nieuw-Vennep, The Netherlands) as a means to interact with a virtual environment. Training with this system led to improvement of function (measured by upper limb section of Fugl-Meyer scale) in chronic stroke subjects [15]. The Driver's SEAT project incorporated a constraint-induced therapy paradigm into a gaming interface (simulation of steering a car) to increase subject motivation and promote coordinated bilateral movement in the upper limbs [16].

Feedback distortion, a growing unconventional therapy paradigm, refers to the use of intentionally inaccurate feedback. In recent studies in the labs of Patton [17, 18] and O'Malley [19], subjects were asked to use a robotic manipulator to move a round spot to a visual target presented to them on a monitor. Since the robot and the subject's hand were obscured, the only visual feedback was from the moving spot on the monitor. Three target points were placed radially from the starting point and subjects were asked to move to a target when it was highlighted and then go back to the starting point. In this repetitive movement training in a virtual environment, subjects showed faster learning when the position of the point was shown further away from the actual position of it by augmenting the initial trajectory error (measured as deviation from the straight line between starting point and target). Subjects also showed the ability to adapt to a visual distortion (implemented by rotating the subject's actual hand position around the starting point) without a visual error augmentation. This type of training was shown to lead to persistent functional changes with hemiparetic post-stroke subjects [17]. A more detailed discussion of this kind of training is presented in Chapter 2 of this thesis.

1.3 Engagement in Therapy: Importance and Contributing Factors

Therapy by nature is hard work. During therapy exercises, in order to regain motor function, patients are required to push themselves beyond their abilities. Therapy regimens are based on our understanding of neuroscience and motor learning principles. To stimulate patients' brain neuroplasticity and reorganization of motor pathways, therapy exercises comprise of numerous repetitions of simple physical tasks. This necessary repetitive nature of therapy exercises makes the hard work of therapy boring, which can negatively affect the patient's motivation to continue therapy [20, 21].

Motivation of the stroke survivor to continue and stay engaged in therapy exercises is clearly a contributing factor to the success of therapy programs. In fact, a qualitative study has shown that patients with low motivation believe they do not have the strength to complete exercise tasks, which in turn will lead to therapy abandonment [22]. Additionally, depression is a common symptom post-stroke and studies have shown that depression negatively affects a patient's self-efficacy for sustained therapy [23, 24], emphasizing the need to find ways of improving patients' engagement during rehabilitation exercises. Sustaining motivation and engagement becomes even more critical as therapy sessions shift from being therapist-based to robot-based; a therapist can observe a patient's physical and emotional condition and shape the exercise task

accordingly, while giving positive and motivating verbal feedback. However, programming this into a robot is a difficult challenge.

To address this issue, robotic therapy research is currently being extended to make it more engaging. For example, virtual reality-based interfaces, video games and immersive experiences increase the patient's engagement in the therapy process. This method, also known as implicit learning, has been proven to be more effective than the conventional paradigm of explicit learning [10, 25].

By utilizing new neurorehabilitation strategies such as error augmentation, the use of robotics in stroke therapy has improved the effectiveness of therapy sessions. Additionally, robot-assisted therapy has the potential of increasing task engagement via the use of aesthetics of virtual reality environments. However, robotic therapy regimens are still completely dependent on the presence of a therapist to set up appropriate exercise protocols based on each patient's strength, performance and emotional state.

Although therapy exercises are difficult, research in psychology and game design suggests that this physical challenge can be used to make therapy more engaging. A large body of research in the fields of psychology and game design has been dedicated to the study of the process of engagement in a task. This knowledge is widely used in the design of games in order to increase gamers' engagement and satisfaction with games, and thus, sell more copies of each game.

Csikszentmihalyi's theory of "flow" explains the mechanics of engagement in a task in addition to what makes a task engaging [26]. Flow is defined in positive psychology literature as the satisfying feeling of heightened functioning in a task with full concentration to finish it. A task that is capable of delivering a flow state has clear goals, as well as a clear feedback system to help the person involved in the task with refining his/her perception of the task challenge and his/her skill.

According to Russell's "Circumplex Model of Affect" [27], every emotional state can be stated in a linear combination of arousal (intensity of the feeling, i.e., low or high) and valence (being positive or negative) as shown by Figure 1.2. A person in the flow condition experiences a state of full attention (high arousal) and joyful immersion (positive valence) in the task at hand. Based on this, flow as a positive affective state is the opposite of several negative affective states such as anxiety (high arousal, negative valence) and apathy (low arousal, neutral valence). These

negative affective states during a task are functions of the task's challenge level and the level of skill of the person involved in the task [28]. For example, individuals involved in a task get bored (low arousal, negative valence) when their skill sets are superior to the skills required to fulfill the task. In this case, the individual can be moved out of the boredom state and into the state of flow by increasing the challenge of the task. This will challenge the person at a desirable level, allowing for acquisition of new skills in addition to practice of old skills.

Guadagnoli and Lee [29] have proposed the "Challenge Point" framework to conceptualize the effects of different practice conditions (i.e., task difficulty) from a motor learning point of view. By looking at the combination of task difficulty and performance instead of task difficulty and skill level, their framework qualitatively defines optimal challenge points. Individuals practicing at these optimal challenge points (i.e., an individual's desirable difficulty) will have the highest potential for learning. It can be argued that at these optimal challenge points, the individual is properly challenged and will learn new skills with a sustainable level of engagement in the practice session. This can be an equivalent to the state of flow.

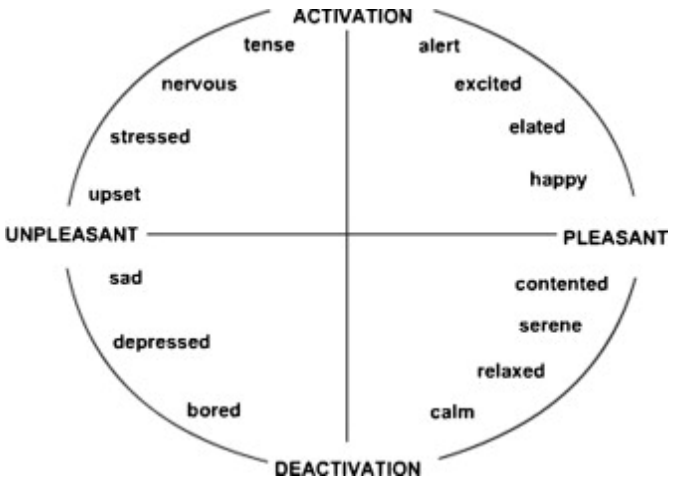


Figure 1.2: Russell's "Circumplex Model of Affect"

1.4 Purpose and Overview of This Thesis

While the main goal of physical therapy is to accelerate stroke survivors' motor function recovery (i.e., motor learning), a therapy regimen cannot be prescribed solely based on a patient's motor abilities; patients' engagement in therapy also plays a vital role in success of stroke rehabilitation programs. According to the flow theory and the challenge point framework, practicing at one's desirable difficulties ensures both learning of new skills and engagement during each exercise session. Furthermore, these desirable difficulties are dependent on one's

performance and affective state during that task. Seeing one's improvement in an engaging task will increase one's self-efficacy with the task, and in the case of stroke therapy, will lead to a lower chance of therapy abandonment and faster recovery.

Implementation of the idea of desirable difficulty relies on two main factors: having an exercise with several different difficulty levels, and offering the exercise to individuals at their desirable difficulty. One of the advantages of robot-assisted therapies is that parameters of a robotic exercise can be easily and quantitatively altered to manipulate exercise challenge. This suggests that the idea of desirable difficulties can be fit into robot-assisted therapy regimens. In that case, the main challenge will be to enable a robotic system to predict the user's desirable difficulty during exercise. As mentioned before, these desirable difficulties depend on the user's performance and affective state during the task. While robots provide the ability to quantify the user's performance, assessing the user's affective state requires more effort and can only be done indirectly.

With the well-established literature on the relationship between emotional state and physiological signals, in the recent years it has been suggested that the engagement level of robotic rehabilitation can be increased by monitoring the patient's physiology during therapy. Real-time analysis of physiological signals, such as electrocardiography, respiration rate, skin conductance and breathing rate, is one way to characterize different aspects of affect. Kulić and Croft developed methods to estimate human affective state in real time using interpretation of physiological signals in a two-dimensional valance-arousal representation [30] using Hidden Markov Models [31] and a Fuzzy Inference Engine [32]. In a recent study, Pan et al. investigated the viability of a physiologically-triggered bookmarking paradigm [33]. In this study, orienting responses (derived through monitoring of the subject's galvanic skin response) to a disruption of attention were used to bookmark electronic media with 84% success, so subjects could resume listening to the media after attending to the interruption. Zoghbi et al. developed an explicit method of real time affective state reporting [34]. In this method, the subjects were asked to express their emotions via an in-house developed hand-held joystick called Affective State Reporting Device (ASRD). A more detailed discussion of studies involving affective computing using physiological signals is presented in Chapter 3.

A possible issue with using physiological signals in rehabilitation can be that such signals are determined by both a person's psychological state and physical activity. However, investigations

by Munih [35, 36] suggest that physiological signals can reliably describe a person's psychological state even in the presence of physical activity, and thus they can be useful in designing a bio-cooperative exercise system.

A good example of such a system is demonstrated by the work of Liu et al. in using psychophysiological signals in a closed-loop exercise system [37]. A robot-based basketball task was designed with the speed of the basket's motion in 3D space (the basket is placed on the robot's end-effector) as the basketball throwing task difficulty. In this study, the robotic coach (i.e., the robot) can change task difficulty based on anxiety (changes aimed to lower the anxiety level) or performance (changes aimed to increase the performance level) of the human player. This study showed that determining challenge level in a human-robot interaction task based on a human's affective state leads to higher performance improvement compared to setting the challenge based only on the person's performance.

This thesis investigated the feasibility of implementing the idea of desirable difficulties in a robot-assisted motor rehabilitation exercise to improve users' engagement in the exercise. This design process was split into three main steps: 1) design of an effective (i.e., leading to motor learning) robotic exercise with meaningful and distinguishable levels of difficulty, 2) development of a method to predict the user's desirable difficulty during the robotic exercise, and 3) evaluation of the effects of adjusting the exercise based on the user's desirable difficulties on the user's perception of the exercise. The results of each step are presented in the following chapters.

Chapter 2 investigates the ways of adding meaningful levels of challenge into a robot-assisted reaching task. Mussa-Ivaldi, Patton, and O'Malley [17-19] have investigated the use of error augmentation (versus teaching via reducing a patient's trajectory error in the conventional hand-over-hand guidance) in a robotic exercise for reaching motion. Their studies have shown that practice with error augmentation leads to faster improvements in motor function. Building on these findings, in a human user study, this thesis investigated the possibility of using different levels of error amplification (EA) as different difficulty levels of reaching motion exercise (i.e., EAs as effective and distinguishable difficulty levels). The results, presented in Chapter 2, showed that while EAs are an effective way of improving the speed of motor adaptation, study participants reported different levels of perceived challenge for each EA level.

Chapter 3 presents the development of a method to predict the direction of change to reach a user's desirable difficulty during the robot-assisted reaching exercise. Participants' physiological data were collected and their motor performance during reaching exercise with different levels of EA was calculated. After practicing with each EA, participants reported the direction of change to reach their desirable level of difficulty (i.e., whether they want an easier or a harder *next* trial). The efficiency of three different machine learning methods were investigated to be able to predict a user's challenge preference based on different combinations of physiological data and motor performance attributes. The choice to include physiological data in this process was inspired by Liu's robotic basketball study [37] (similar to Liu's study, in the studies presented in this thesis, participants performed a hand-eye coordination task with different difficulties in a distorted environment).

In a final study presented in Chapter 4, the effects of practicing under the desirable difficulty condition were studied. In this study, the control group received EAs in predefined random order, while the experimental group received EAs based on the predictions of the trained machine learning algorithm. A post-experiment questionnaire was used to assess user preference between these two conditions. The experimental group reported statistically significant higher scores on the metrics that assessed satisfaction, attentiveness, perceived performance, and willingness to continue the task. Additionally, this study examined fidelity of our desirable difficulty prediction algorithm based on accuracy of its predictions for the experiment group.

Chapter 5 discusses the implications of this work in the field of robot-assisted therapy, with a focus on summarizing the findings of this work, as well as covering recommendations for future work.

2 Study I: Error Amplification to Promote Motor Adaptation, Alter Affective States, and Create Differences in Challenge

The first step to implement a system capable of delivering therapy exercises at users' desirable difficulties is presented in this chapter. Focusing on reaching motions that only involve the use of shoulder and elbow joints, reaching trajectory error amplification via visual or/and force feedback was proposed as a way to increase the efficiency (i.e., accelerating motor adaptation and adding several challenge levels) of a simple massed-practice, point-to-point robotic reaching exercise.

In the study presented in this chapter, 10 healthy participants completed blocks of reaching tasks similar to those of [17-19] with different levels and methods of error amplification (i.e., low gain vs. high gain and visual vs. visual plus force feedback), and reported their satisfaction, attentiveness and dominance over the robot and the task during each block. Section 2.2 presents a literature review of reaching motion planning in stroke and healthy populations and available ways of robot-assisted rehabilitation of those motions. Details of the experimental and analytical methods used in this study are presented in Section 2.3. Section 2.4 covers results, and Sections 2.5 and 2.6 discuss the results and provide conclusions of the study.

2.1 Introduction

2.1.1 Reaching Motions in Healthy and Stroke Populations

The use of wrist, elbow, and shoulder are involved in reaching to targets within one's arm length; however, reaching to targets beyond arm's length requires recruiting hip and trunk motions in addition to moving wrist, elbow, and shoulder joints [38]. Healthy individuals use these joints in coordination with each other to accurately place their hands over a target. The excess of degrees of freedom (DOF) in the human arm (i.e., 7 DOF) enables the central nervous system to utilize an infinite number of paths in reaching to a target. Nevertheless, healthy individuals use largely the same joint motions during reaching to targets in a similar reaching condition (i.e., holding the start and end position and hand orientation constant) [38, 39].

In the healthy population, the central nervous system controls the motion of hand by transforming sensory signals into hand trajectories, and subsequently into matching joint trajectories, required muscle torque, and finally into a pattern for activating muscles [40]. Hogan

et al. [41] characterized the normal reaching motion to targets within one's arm length. These motions follow a bell-shaped velocity profile, in which a peak velocity occurs between the start and end point. The position of this peak velocity is an indication of one's strategy in increasing accuracy versus speed of reaching.

A straight or slightly curved hand path is generally followed when reaching for targets within arm's length [41, 42]. To assess performance of individuals in performing reaching tasks, different measures for the straightness of the hand path have been proposed. Bastian has used the ratio of the actual hand path length to the direct path between the start and end points of reaching [43]. Research groups led by Patton and O'Malley [18, 19] have used the maximum deviation of the actual hand path from the direct path as a measure of straightness. Deviations from these straight paths, a predominant condition in stroke survivors, are caused by spasticity, decreased range of motion, reduced elbow-shoulder coordination, and muscle weakness. Altering this condition in stroke survivors and enabling them to use their elbow and shoulder in performing reaching tasks is usually one of the goals in stroke physical therapy.

2.1.2 Use of Robotics in Motor Rehabilitation of Reaching Motions

Therapy robotics is mainly focused on providing a means of delivering the physical part of the therapy. Most of the available robots for therapy can be used in either passive, assistive, or resistive modes [10]. To rehabilitate the reaching motion in a passive exercise session, a patient's paretic arm is moved by the robot without requiring any muscle input by the patient. This is a proper exercise for patients with low range of motion and higher degrees of muscle weakness. As a patient improves, training changes from passive to assistive. In assistive training, the robotic system provides extra motion to help patients complete a task they themselves have initiated. Finally, in resistive training, the robot resists the motion of the arm, requiring patients to increase their physical effort. Different studies using different robotic systems have shown the effectiveness of training with robotic systems in motor rehabilitation after stroke (see Section 1.2).

With the advancements in haptics technology and the development of immersive virtual-reality interfaces, robotic therapy regimens that supersede therapists' physical exercises are being developed.The MIME system and the MIT-MANUS are two of the well-known robotic exercises that can be used to augment conventional rehabilitation techniques either independently or in

conjunction with virtual reality training. The MIME system utilizes a bimanual training approach and the MIT-MANUS provides a high dose unimanual massed-practice approach.

To implement bilateral training, two robotic arms are used in the MIME system [13]. Each of the participant's hands is strapped to one of the robotic arms via a splint. One of the robotic arms follows the non-paretic arm to capture its trajectory in real-time, while the other robotic arm guides the paretic arm by mirroring the motion of the non-paretic arm. This guidance can be either passive or assistive. On the other hand, the MIT-Manus robotic exercise is a unimanual reaching exercise involving the paretic arm of participants [11], delivering a massed practice exercise. For the exercise task, the robot is programmed to control the motion of a cursor on a monitor. During an exercise session, participants are asked to move the cursor to presented target points on the screen by manipulating the robot's end-effector. Similar to the MIME system, the MIT-Manus robot is used to guide the participant's motion either actively or passively.

While a human therapist initiates the motor learning by showing the correct trajectory of motion to the participant (reducing the trajectory deviation), in robotic therapy exercises, a promising teaching strategy is to emphasize the amount of trajectory deviation (augmenting the trajectory deviation). Research in both artificial intelligence (e.g., Neural Network design) and motor learning proposes an error-driven process that supports learning [44, 45], suggesting that error augmentation can be used to maximize the effect of massed practice.

Use of feedback distortion and error augmentation in motor learning and adaptation has been extensively studied for both fine and gross motor functions. In a study focusing on the learning of fine motor movements (i.e., grasping), Matsuoka et al. demonstrated that distortion of the visual feedback improves the pinching movement patterns of healthy participants [46]. Building on this study, the same group showed that such distortion of visual feedback can also be used to improve grasping function in the post-stroke population [47].

Wei et al. compared the effects of two error augmentation methods – visual error offsetting and visual error amplification – in learning of reaching motion as a gross motor function [18]. Introducing maximum deviation from the ideal straight path in a reaching motion as a performance measure, they studied speed and amount of adaptation to a rotational visual field in a robotic point-to-point reaching exercise. This study showed that these methods significantly increase the speed of motor adaptation in healthy subjects and improves the speed of motor

learning in stroke survivors. With the same task and environment, Celik et al. compared progressive visual error offsetting, their novel error augmentation method, and the two methods used by Wei [19]. They used target hit time as a second measure of performance, in addition to the maximum lateral deviation in the reaching trajectory, and modified definitions for "speed" and "amount" of learning to further supplement the findings of Wei.

In addition to these visual distortion methods, a physical reaching environment can also be distorted by force feedback. Patton et al. showed that training the same point-to-point task, where force feedback in the direction of trajectory errors was provided via the robot's end-effector, facilitated a higher rate of motor adaptation to a force field in a healthy population and a higher rate of reaching trajectory improvement in a clinical population [17]. Note that this kind of force feedback in the direction of trajectory deviation is different than the force feedback in resistive training that is opposite the direction of motion. His study also demonstrated that after removing the force field (i.e. reaching in a non-distorted environment) the clinical population retains the functional improvement gained from practicing with the force feedback. However, this was not true for the healthy population, and the effects of practicing with the force feedback were washed out in the follow-up study. Moreover, Patton et al. showed that practicing with error augmentation, reinforced by the therapist's verbal feedback, leads to a higher range of motion and elicits performance improvement as measured by the arm motor section of the Fugel-Meyer (AMFM) scores for post-stroke participants (using $p < 0.1$) [48].

2.1.3 Purpose and Overview of the Study

The repetitive nature of massed practice of simple movements (like robot-assisted reaching to visual targets) to achieve any measurable functional gain can be considered a downside to the aforementioned error-augmenting exercises proposed by prior work. While use of feedback distortion can lead to faster improvement of reaching movements, participants' engagement within such exercises, as well as the level of perceived task challenge, has not been studied.

This study investigates the effects of combining force feedback and visual error amplification on adaptation to a visually distorted environment, similar to [18, 19], in performing robot-assisted point-to-point reaching motions in a healthy population. The study of retention or wash-out of this kind of adaptation was not performed and thus, is not included in this chapter. Moreover, as error amplification makes the repetitive reaching task more challenging, it is expected to see a higher engagement during reaching with error amplification. Based on this, this study aimed to

measure the relative effectiveness of each error amplification method in promoting motor adaptation to a visually distorted environment and increasing participants' engagement in the task, in comparison with the control condition (i.e., no error amplification).

A real-time controller for a robotic manipulandum was modified to study the effects of different feedback error amplification methods on a participant's upper-limb motor learning and affect during a point-to-point reaching exercise (Appendix A). We conducted a study in which the reaching environment was visually distorted by implementing a 30° rotation between the coordinate systems of the robot's end-effector and the visual display. Feedback error amplification was provided to participants as they trained to learn reaching within the visually rotated environment. Error amplification was provided either visually or through both visual means and force feedback, each method with two different amplification gains. Participants' performance (i.e., trajectory deviation) was used to study the speed and amount of adaptation promoted by each error amplification method. Self-reports to a questionnaire were used to study participants' affective state changes caused by each error amplification method.

2.2 Methods

2.2.1 Research Ethics and Study Participants

Ten healthy participants participated in this study: five males and five females, one left-handed, with an age range of 19-27. Participants provided informed consent as approved by the Clinical Research Ethics Board (CREB) of the University of British Columbia. To make sure participants were cognitively intact and free of any significant neurological impairment, all participants were required to score higher than 24 on the Folstein Mini-Mental State Test. All the participants had normal or corrected vision to eliminate any correlation between task performance and vision in the hand-eye coordination task used in this study.

2.2.2 Hardware

A five-bar robotic manipulandum previously developed at UBC was used in this study [49]. Figure 2.1 shows a participant interacting with the robotic manipulandum by holding onto its end-effector. The manipulandum's controller runs at 1 kHz and its end-effector sweeps a two DOF horizontal working area of approximately 50 cm × 35 cm. Two motors (Parker-Compumotor Dynaserv DR1060B) located at the base (robot's shoulders) actuate the robot, and the two "elbow" joints are passive and not actuated. The robot's end-effector is a handle

instrumented with a 6-axis force-torque sensor (ATI Industrial Automation Inc. "Mini" sensor). In the schematic drawing of Figure 2.1, the active and passive joints as well as the end-effector are illustrated by yellow dots and the blue lines show the robot's links.

In the control architecture of the robot only forces in the horizontal plane (i.e., the robot's work space) are used. Encoders integrated with the motors supply position feedback. Using TargetDisplay (MathWorks Inc.), the position of the end-effector is visually rendered on a flat screen monitor as a moving dot at 20 Hz. In this study, the robot's original controller was modified and re-programmed to deliver reaching exercises with visual distortion and error amplification. More details are given in Appendix A.

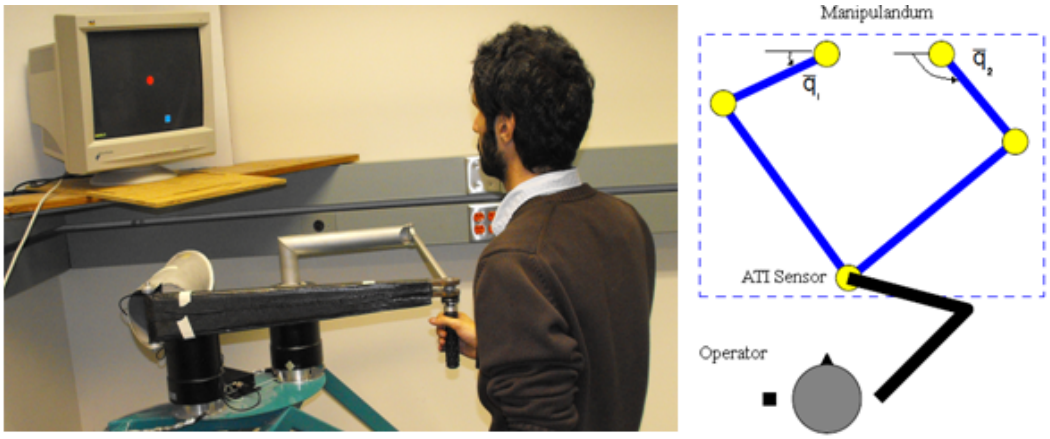


Figure 2.1: Left: a participant interacts with the robotic manipulandum. Right: a schematic top view of the reaching environment.

2.2.3 The Reaching Task, Visual Distortion, and Error Amplification Levels

To complete the robotic reaching task, participants were instructed to move the movable dot (i.e., the cursor) to visual targets presented on the monitor by manipulating the robot handle with their non-dominant hand. In this massed-practice training, the 17" monitor was mounted in front of the participant, with a cover occluding the reaching space. The goal of this occlusion was to make sure that the only visual feedback available to the participants was through the monitor and not by observing one's hand during the reaching exercise (Figure 2.2). Participants were told that the moving dot exactly follows their hand position and shows the actual position of the handle. Three targets were placed radially, 120° apart, at a constant radius from the start position (i.e., middle of the screen). When a target was highlighted, participants had to move the robot handle to place the moving dot over the target, and then move back to the start position. Each reaching motion covers approximately 17 centimetres. Each three consecutive reaches is called a *cycle*, in which the three targets appear in random order.

In order to ensure all participants completed reaching motions in a relatively consistent manner (i.e., speed and accuracy of reaching), reaching duration was fixed at 0.5 seconds. This means that participants were asked to complete each reaching motion in 0.5 seconds. Each target was visually presented as a filled blue square surrounded by a green square (Figure 2.3). As soon as a participant started to move the cursor toward the blue target square, the green square would start to shrink at a constant rate to disappear into the blue square after 0.5 seconds. In order to assist the participants with timing of their reaching motions, they were instructed to reach to the blue square in a single motion and stop on the target just as the green square disappeared. They were asked to remain at the location of achieved target and wait for the next target.

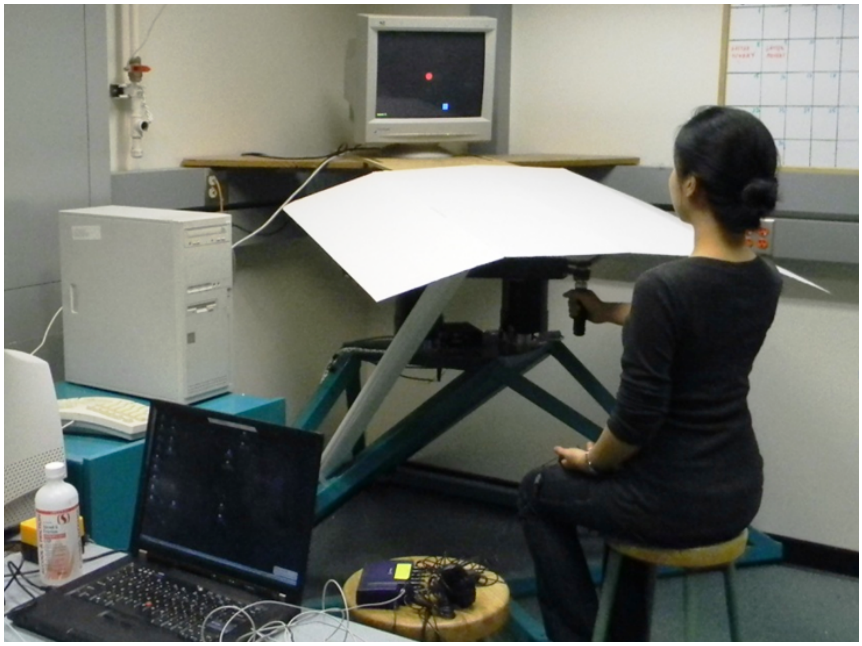


Figure 2.2: A participant completes a reaching task while her hand is occluded by a cover.

To make the task more entertaining to the participants, 1, 2, or 3 points were awarded for each reaching motion based on the participant's motion timing accuracy: 3 points for a reaching time within 10 milliseconds of the target time of 0.5 seconds (i.e., reaching time between 0.49-0.51 seconds), 2 points for being within 50 milliseconds of the target time, and 1 point for being within 100 milliseconds of the target time. A score counter displayed on the screen showed the session's total score, and participants were encouraged to try to achieve a high score (i.e., more consistent reaching time, which means a more consistent reaching trajectory planning). Figure 2.3 shows a screenshot of the manipulandum's monitor: the red circle is the cursor and participants were asked to move its centre into the blue square (i.e., target point) by using the green square to time the motion. A counter shows the session's cumulative score.

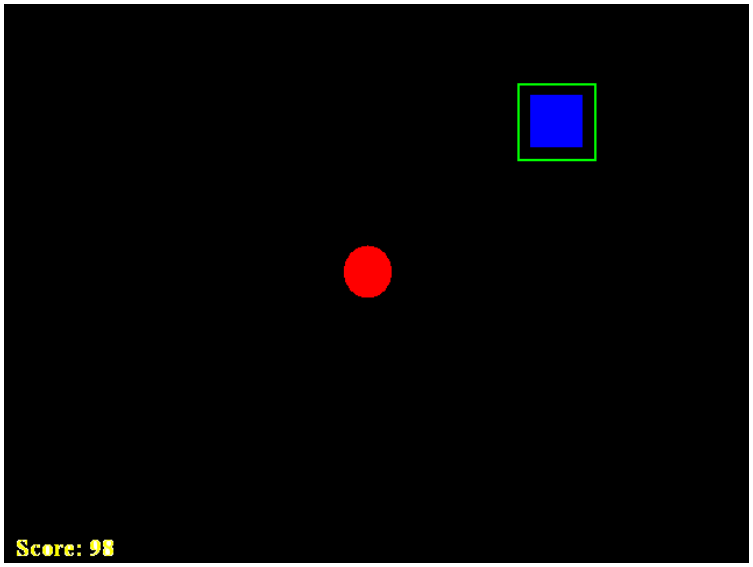


Figure 2.3: A print-screen of the manipulandum's monitor to show the relative sizes of the cursor and target point. These sizes were kept constant for all participants.

The goal of this thesis is to show viability of the "desirable difficulties" method in addressing the issue of engagement in a repetitive task. As stroke survivors are harder to recruit and studies with them are more time consuming, they were deemed a less desirable population to test during methodology validation studies such as this one. Thus, healthy participants were recruited instead.

After several trials of the robotic reaching task, healthy participants generally demonstrated an ability to complete the task with minimum deviations from the target time and the ideal straight path. Studying motor adaptation to visuomotor distortions in healthy participants in methodology validation studies is proposed as an alternative to studying motor learning in stroke survivors [18, 19]. In that case, visuomotor distortions can be considered as a proxy for stroke survivors' damaged neural pathways and muscle synergies (for instance completing reaching tasks in a force field can be analogous to muscle tightness following stroke).

Motor adaptation can be considered as a special case of motor learning that can happen over a short period of time, during which a healthy individual learns to carry out a task in a distorted environment. The main difference of motor adaptation in a healthy population and motor learning in stroke survivors is in retention of the gained functional improvements from training. For able-bodied participants, adaptation is quickly followed by a wash-out of the distortion effects. However, a study in Patton's lab showed that for hemiparetic post-stroke participants, this type of training can lead to persistent functional changes, evidencing adult neural plasticity.

Adaptation in healthy participants occurs in a short period of time (a one-hour training session). The rapidity of this process allows for the characterization of adaptation by measures such as amount and speed of improvement, and performance after training session (see Section 2.3.2). On the other hand, motor learning in the stroke population is a very slow process. Performance during the motor rehabilitation period is measured on standard clinical scales such as the Fugl-Meyer Assessment or the Wolf Motor Function Test. Their scores change over weeks of practice (compared to performance scores for motor adaptation that can be calculated after each training block).

To increase participants' trajectory deviation and to initiate motor adaptation a visual distortion was implemented. To implement this visual distortion, actual hand movement (handle position) was rotated 30°, and then the rotated position was presented as the moving dot on the monitor. This means that a hand motion toward the right of the screen to achieve a target would be presented by a cursor path that is 30° off toward the bottom of the screen (Figure 2.4). Participants were told that the cursor follows their hand position, thus visual distortion acts as a deception.

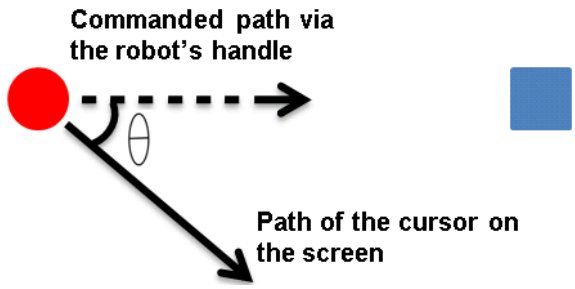


Figure 2.4: In a visually distorted environment cursor does not follow the actual hand paths. θ=30° is the amount of distortion.

As expected, during the first few reaching trials in the visually distorted environment (i.e., first exposure) participants were not able to follow the straight path and their reaching path was curved. Participants trained in the visually distorted environment to adapt to the distortion and be able to follow a straight reaching path between the start and end position. Reaching without error amplification was used as a control condition in the analysis of motor adaptation to the visually distorted environment.

At each point on the reaching path, the deviation vector from the straight path is called reaching trajectory deviation (Figure 2.5). In training with visual error amplification, an augmented error vector (by a gain of α) was added to the vector of hand path in the distorted environment. Then

this error-amplified hand trajectory was used to update the cursor position. In addition to the visual error amplification, force feedback can also be utilized to amplify the trajectory deviation. In that case, the deviation vector was used to calculate an augmentation force vector.

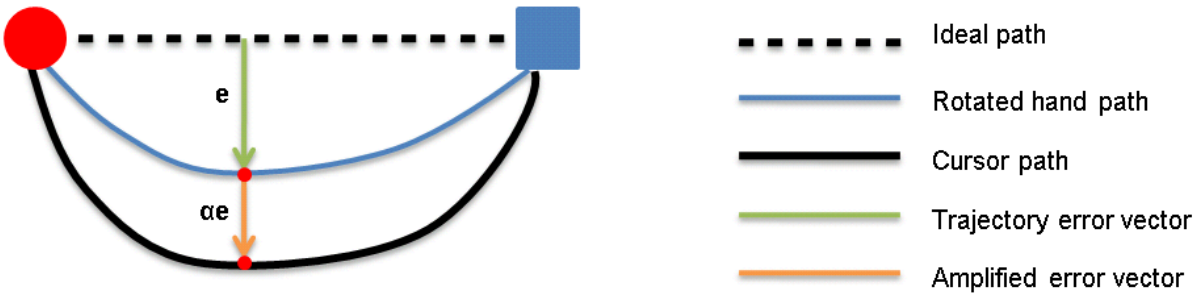


Figure 2.5: The deviation from the straight path between the start and end points of reaching motion was used as an deviation vector and was amplified via visual means and force feedback.

Visual error amplification was implemented as described above and similar to the methods presented in work of Celik et al. in [19]. Two amplification gains of α were used in this study: a low gain of 0.30 and a higher gain of 0.65. In this error amplification condition, vector summation of rotated hand position and amplified error vectors was used to calculate cursor location:

$$x_{\text{cursor}} = x_{\text{hand, rotated}} + (\text{visual EA gain}) \times e \tag{2.1}$$

In addition to the visual error amplification, force feedback was provided to the participants based on their instantaneous trajectory deviation vector. The error amplification force vector, perpendicular to the vector from starting point to the visually presented target, was calculated using the following equation:

$$F_{\text{Feedback}} = (\text{Force Feedback EA gain}) \times e \tag{2.2}$$

This force was exerted onto the participant's hand via the robot's end-effector during reaching motions. Force feedback error amplification gains were designed to map the trajectory deviations to force ranges of 0-5 N and 0-8 N, for low gain and high gain, respectively.

The efficiency of force feedback error amplification in improving the recovery speed has been validated in the stroke population [17, 48]. The effects of different visual error augmentation methods have been extensively studied by the O'Malley and Patton research teams [17-19]. Building on these studies, one of the contributions of this thesis is implementing and assessing

the *combination* of the two ways of implementing error amplification (i.e., visual and force feedback) on motor adaptation in the healthy population.

2.2.4 Study Protocol

After the participants signed the consent form, the Folstein Mini-Mental State Test was administered (Appendix C). All of the participants scored higher than 24 on this test, indicating that they are free from major neurological dysfunctions and were able to complete the study. After this, the robot, the reaching task, and the scoring scheme used to reflect the timing of each reaching trial were explained to the participants. Before starting the study, the participants were seated in front of the robot, facing the robot with their trunk parallel to the robot's monitor at an approximate distance of 30 cm from the robot's end-effector. The participants were asked to sit at this position to make sure the effects of visual distortion were the same for all of them (sitting non-parallel to the monitor can be viewed as adding/decreasing the visual distortion). Also, the approximate distance of 30 cm ensured that the participants did not need to go beyond their arm length to reach for the visually presented targets. Participants were instructed to only use their shoulder and elbow to move the robot's handle and avoid using their trunks to control the end-effector motion more accurately. On average, each experiment session took 90 minutes to complete.

The experimental protocol (Figure 2.6) was designed as six exercise blocks. In the first block, participants performed 14 cycles of reaching tasks without any rotational distortion (i.e., plain motions). Each cycle comprises of one reaching motion from the center point of the screen to each of the three visual targets in a predefined random order. This block aimed to get participants familiar and comfortable with the robotic reaching task. Following the familiarization block, there were five training blocks. Participants were given a rest period upon request at any time, in addition to one-minute rest periods between the training blocks.

Within each of the training blocks, participants practiced 10 cycles of plain motion (de-adaptation), and 13 cycles of reaching with one of the designed challenges "within" the visually distorted environment (exercise with challenge). Visual distortion was implemented as either a −30° or a +30° rotational field (randomly chosen) in each of the exercises with challenge sub-blocks. According to prior work, adaptation to a visual distortion only depends on the degree of rotation (30° or 45°) and does not depend on the direction of the rotation (-30° or +30°) [17]. This change in the direction of rotation was to reduce the possibility for participants to figure out

what is causing their performance to decrease at the beginning of each exercise with challenge. This further ensures that visual distortion remains a deception to the participants.

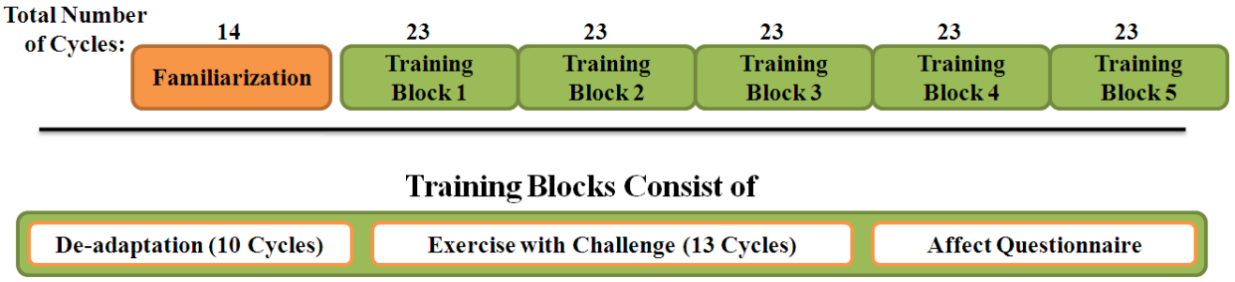


Figure 2.6: Top: Detailed break-out of the experiment protocol. Participants were given this information prior to the start of the experiment. Bottom: A break-out of a training block.

The five challenges (one per training block) utilize different methods of error amplification to promote motor learning. Each participant practiced the challenges in random order. The de-adaptation cycles at the beginning of each training block are designed to washout the learning effects (i.e., adaptation to the visual distortion) of the previous training block, so adaptation does not carry over from one challenge to the next challenge. Pilot studies showed that participants become de-adapted after 7.3±1.1 cycles. Using a 95% confidence interval, the number of required de-adaptation cycles was calculated to be 10.

At the end of each training block, a Self-Assessment Manikin (SAM) affect questionnaire was administered [50], asking the participants to self-report their satisfaction, attentiveness, and level of control over the task and the robot (see Appendix B). Satisfaction scores ranged from 9 for being annoyed to 1 for being pleased. However, in the results section (Section 2.3.4) satisfaction scores are not presented as an inverse item. Attentiveness scores ranged from 9 (stimulated) to 1 (relaxed). Scores for level of control ranged from 9 from being in control of the robot to 1 for being dominated by the robot.

The SAM questionnaire is based on the model for affect proposed by Russell [27, 30]. Russell's "affect grid" is used to express human emotions in two dimensions: valence and arousal. Valence quantifies if a feeling is positive or negative and arousal quantifies how strong that feeling is. In the SAM questionnaire, the words "valence" and "arousal" are changed to two more familiar words: satisfaction and attentiveness. The SAM questionnaire also adds a third dimension, that is, level of control, to the two-dimensional model proposed by Russell. It is more common to use questionnaires that only use the two dimensions of valence and arousal in human-robot interaction studies to extract data about a user's feelings. However, in this thesis

the SAM questionnaire was used to collect data about the participant's perceived level of control, in addition to the participant's satisfaction and attentiveness. This third dimension of affect was used to study the level of perceived challenge in each of the training blocks.

Five conditions of error amplification (EA) were used as challenges in training exercises for adaptation to the rotational field: reaching without EA (control), reaching with low-gain visual EA, reaching with high-gain visual EA, reaching with low-gain visual plus force feedback EA, and reaching with high-gain visual plus force feedback EA. This blocked experimental design was used to assess the effects of each training challenge (i.e., error amplification condition) on motor adaptation and the participant's perception of challenge, as well as the affective state of the participants.

2.3 Analysis and Results

2.3.1 Measure of Reaching Accuracy

In order to study motor adaptation, measures of reaching accuracy (Section 2.3.1) and performance (Section 2.3.2) are introduced. Each reaching cycle comprises of a motion to each of the three targets (i.e., three trials). In each of these trials, the actual hand trajectory is measured by the position of the robot's end-effector. For each of the trials, the maximum absolute deviation of the hand path from the line between start and target point was calculated (i.e., the maximum deviation (e) value in Figure 2.5). The average of this value in a cycle (i.e., mean of a cycle's maximum deviations) was assigned as the measure of reaching accuracy for that cycle.

2.3.2 Measures of Motor Adaptation and Performance

The data from the "exercise with challenge" cycles (Figure 2.6, bottom) were used in analyzing different aspects of adaptation. In each of the challenge exercises, participants practice the reaching task in the visually distorted environment while receiving a randomly chosen error amplification method. Figure 2.7 shows reaching paths of participant 4 in the first challenge exercise. Dashed red lines are the participant's hand path in the first cycle of reaching motions in this challenge exercise (i.e., first exposure to the distortion). Blue solid lines are indications of the participant's hand path in the last two reaching cycles of the challenge exercise (i.e., adapted motions).

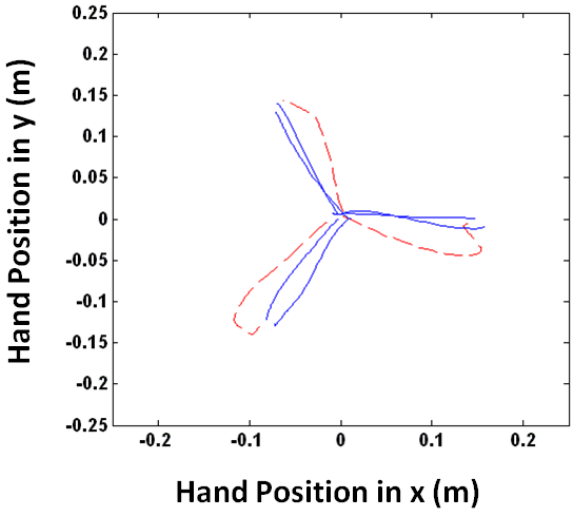


Figure 2.7: Hand path of participant #4 during one of the challenge exercises: the dashed red line shows the reaching path during the first reaching cycle, and the blue solid lines show those of the last two reaching cycles.

Figure 2.7 qualitatively shows the adaptation to the rotational field: at the first exposure to the visual distortion, the participant shows relatively large deviations from the ideal hand path. With practicing the reaching task within the visually distorted environment, the participant adapts to the distortion and adjusts his hand path to go back to the straight line between the start and end points of reaching. This decline in the deviation from the ideal path can be a measure of motor adaptation.

For each of the five challenge conditions (i.e., error amplification), each participant completes thirteen cycles of reaching to the three targets. For each of these thirteen cycles a value for the measure of reaching accuracy (see Section 2.3.1) was calculated. Figure 2.8 shows a plot of reaching accuracy versus reaching cycle number for participant #4 in the first challenge exercise shown in Figure 2.7. As the participant gets more practice in the distorted environment and adapts to it, the average deviation from the straight line decays over the practice time. Practice time is measured by the number of reaching cycles. Wei [18] proposed that an exponential function, such as in Equation 2.3, can be fit to the data.

$$y = ae^{-t/b} + c \tag{2.3}$$

In Equation 2.3, y is the average deviation in cycle and t is the cycle number (numbers between 0-12). Based on this, c will be the convergence value of the reaching accuracy showing the final performance (i.e., the best performance level after training within a specific error amplification condition), b represents the time constant of converging to the c value and can be presented as

the speed of adaptation (higher *b* implies slower learning), and *a* represents the total amount of improvement (i.e., *a* is the amount of decrease in the maximum deviation after practicing within a specific condition). These definitions are visually presented in Figure 2.8. Note that although performance plateaus after cycle number 8, 13 cycles of reaching exercise were performed in order to extend the exercise time without the functional improvement being apparent for the participant. This is a usual condition in therapy sessions that leads to boredom and frustration for patients.

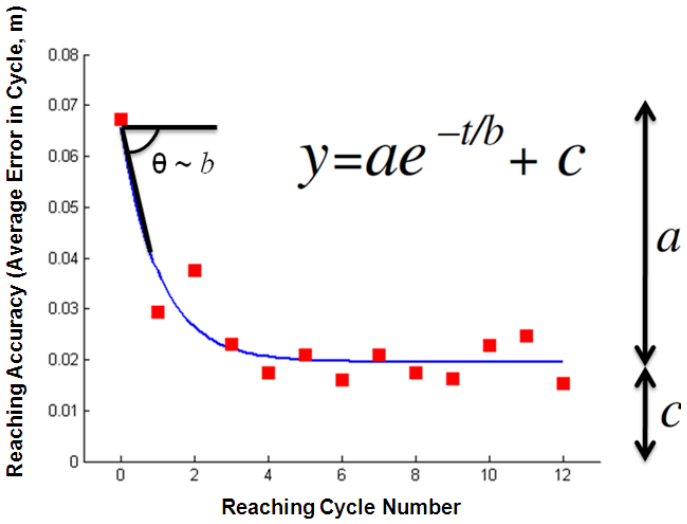


Figure 2.8: Average trajectory deviation in reaching cycles decays exponentially as the participant adapts to the distortion field with practice.

The parameters *a*, *b*, and *c* are used as measures of motor adaptation and performance during a training sub-block with challenge (i.e., reaching within the distorted environment with an error amplification method). As mentioned in the study protocol (Section 2.2.4), each participant completes one training sub-block for each of the five error amplification methods. From the data collected from each participant 5 sets of *a*, *b*, and *c*'s (one per error amplification method) were extracted.

2.3.3 Do Different Error Amplification Levels Lead to Different Levels of Motor Adaptation?

Figure 2.9 provides an example of a qualitative comparison of a participant's (participant #4) adaptation to a visual distortion with practice in different error amplification conditions. For this participant, the visual distortion was implemented by a ±30° rotation between the robot's end-effector coordinates and the display coordinates. As learning is independent of the direction of rotational field, the rotation angle was randomly varied between challenges (either –30° or +30°)

to reinforce deception of the visual distortion. Dashed red lines are the initial hand paths and blue solid lines are paths of learned motions. It is noticeable that the two visual plus force feedback error amplification methods cause higher initial trajectory deviations compared to the other conditions. This implies that the total amount of improvement "*a*" is higher in these two conditions. However, practice in all five conditions leads to adaptation to the distortion, and toward the end of each challenge block the participant was able to follow the ideal path.

In order to investigate the significance of these trends and to quantitatively compare the effects of different error amplification methods on adaptation, a set of *a*, *b*, and *c* as measures of motor adaptation and performance was calculated per participant per exercise with challenge. This provided a set of 10 (i.e., the number of participants) within-subject measurements of *a*, *b*, and *c* for each of the five error amplification methods. These repeated measures were used to statistically compare adaptation under the different error amplification conditions.

The average of the performance metrics *a*, *b*, and *c* and curve fittings are given in Figure 2.10. The following trends can be observed from Figures 2.10 and 2.11: Low-gain visual plus force feedback EA shows the highest amount of improvement (*a*=45 mm), followed by high-gain visual plus force feedback EA, high-gain visual EA, and control. Low-gain visual plus force feedback EA also has the fastest learning rate (i.e., lowest *b*), followed by the other EA types in the same order as the amount of learning. However, this order is reversed for the final performance "*c*". Low-gain visual EA has the poorest learning characteristics.

To compare the effects of different EA conditions on *a*, *b*, and *c*, a within-subjects multivariate ANOVA was performed (Table 2.1). It was found that training with different EA methods does not lead to a significantly different rate of learning and final performance (*b* and *c*). Nevertheless, high-gain visual plus force feedback EA leads to a significantly larger amount of improvement *a*, in comparison with both of the visual EA methods ($p < 0.05$). The results presented in Table 2.1 are based on a multivariate repeated-measures ANOVA and post-hoc analysis with Bonferroni correction. A significance level of $\alpha = 0.05$ was used for all inferential statistics.

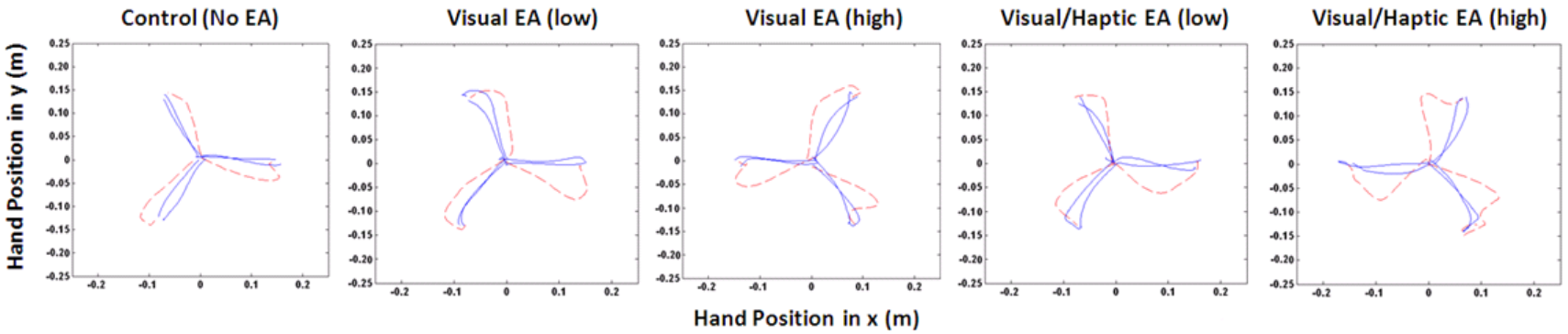


Figure 2.9: reaching path of subject 4 during each of the challenge exercise blocks qualitatively shows the adaptation. The dashed red line shows the paths during the first reaching cycle, while the blue solid lines show the paths during the last two reaching cycles.

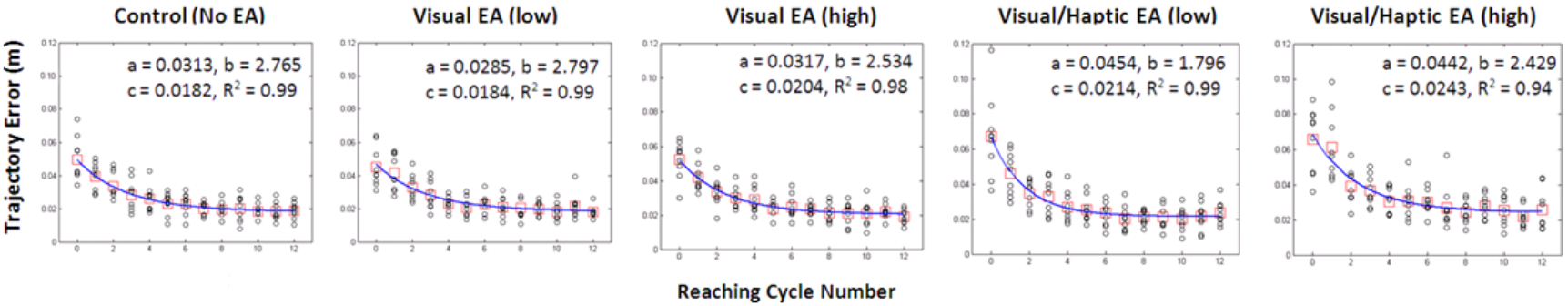


Figure 2.10: Reaching cycle trajectory error (TE) for all the EA methods. Black circles show cycle mean of maximum TE (performance metric) for each subject. Red squares show the average of performance metric for all subjects. Blue curves show the fitted exponential function to the average of performance metric data.

Table 2.1: Repeated-measures ANOVA results comparing motor adaptation metrics across error amplification types.

Measure	ANOVA (EA type)	Post-hoc Analysis
a	$F(1.700, 0.001) = 2.724$, $p = 0.104$	High-Gain Vis&Force and Low-Gain Vis EAs: $p = 0.007$ High-Gain Vis&Force and High-Gain Vis EAs: $p = 0.008$
b	$F(2.601, 2.371) = 0.416$, $p = 0.706$	No significant differences
c	$F(2.940, 6.09E\text{-}5) = 2.221$, $p = 0.110$	No significant differences

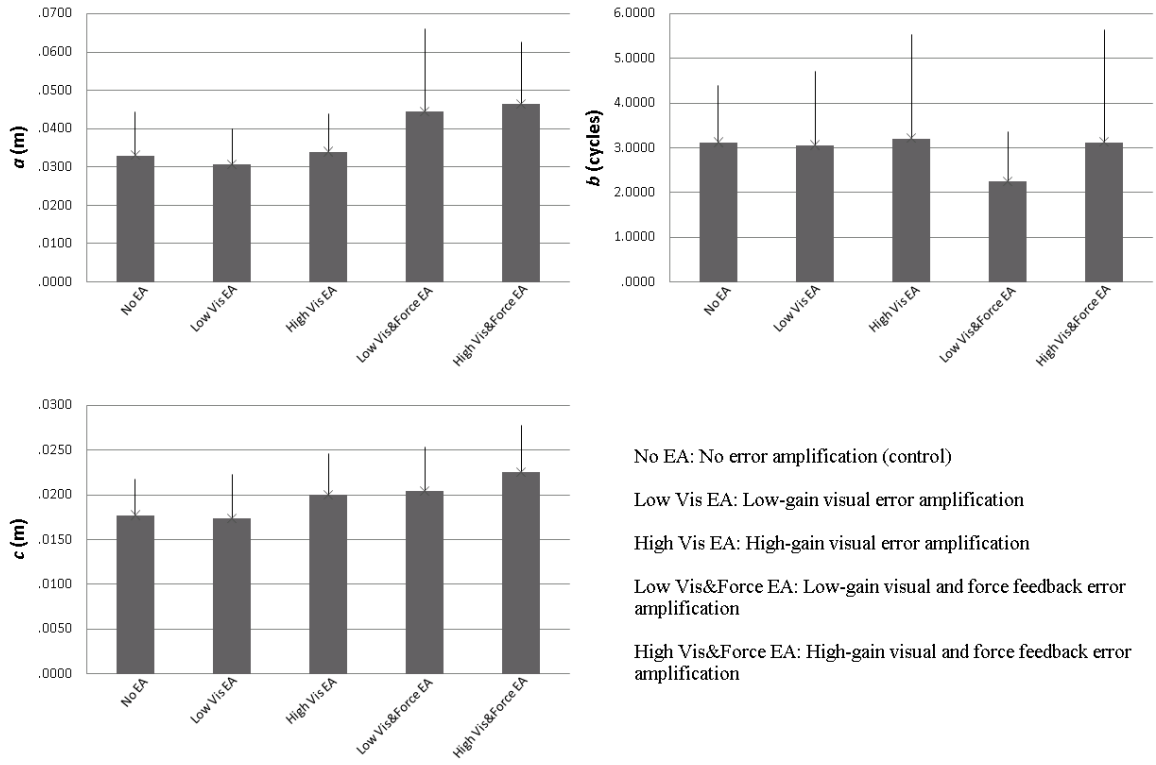


Figure 2.11: Overview of motor adaptation measures demonstrating trends in amount of improvement (*a*), speed of adaptation (*b*), and final performance (*c*) across error amplification types.

2.3.4 Do Different Error Amplification Levels Lead to Different Levels of Task Satisfaction and Attentiveness?

At the end of each exercise with challenge (i.e., practice with different error amplification conditions) participants' self-reports of satisfaction and attentiveness during the period of training were collected. This provided a set of 10 (i.e., number of participants) within-subjects measures of satisfaction and attentiveness for each of the five error amplification methods. These repeated measures were used to statistically compare the effects of exercise with an error amplification method on eliciting different levels of task satisfaction and attentiveness in the

participants (Table 2.2). From control to visual EA to visual and force feedback EA, and from low gain to high gain, the error amplification methods are abbreviated as follows: No EA, Low Vis EA, High Vis EA, Low Vis&Force EA, and High Vis&Force EA.

Table 2.2: Repeated-measures ANOVA results comparing self-reports of satisfaction, attentiveness, and level of control over the task across error amplification types as the within subjects condition. Post-hoc analysis shows significant differences between most of the pairs. Only pairs with non-significant differences are presented in the table. Measures showing significant ANOVA results are indicated with the following suffix: *** $p < 0.001$.

Measure	ANOVA (error amplification type)	Post-hoc Analysis
Satisfaction*** (all pairs sig. except two pairs)	$F(2.301, 46.157) = 43.054$, $p < 0.001$	No EA and Low Vis EA: $p = 0.484$ Low and High Vis&Force EAs: $p = 0.187$
Attentiveness*** (all pairs sig. except two pairs)	$F(3.005, 48.456) = 65.520$, $p < 0.001$	No EA and Low Vis EA: $p = 0.085$ Low and High Vis&Force EAs: $p = 0.107$
Level of control*** (all pairs sig. except one pair)	$F(2.320,47.330) = 44.514$, $p < 0.001$	High Vis EA and Low Vis&Force EA: $p = 1$

Figure 2.12 shows the overall trend in the level of participants' task satisfaction and attentiveness in practicing with different error amplification conditions. The self-reports to the SAM questionnaire (satisfaction was collected in reverse scale, presented in Figure 2.12 as a score) showed an increase in the levels of satisfaction with the task and attentiveness during the task as the EA methods changed from control to visual EA to visual plus force feedback, and from the lower gain to the higher. A within-subjects multivariate ANOVA and post-hoc analysis showed that the means of each of these two affect measures are significantly different between almost all pairs of EA conditions (at highest $p < 0.05$). The exceptions are the following four pairs: 1) level of satisfaction between control and low-gain visual EA, and the two visual plus haptic EA methods, 2) level of attentiveness between control and low-gain visual EA, and the two visual plus force feedback EA methods. The p values for these pairs are given in Table 2.2.

2.3.5 Do Participants Associate Different Error Amplification Levels with Different Challenge Levels?

At the end of each exercise with challenge, participants' self-reports of perceived dominance (i.e., level of control) during the period of training were collected. This provided a set of 10 (i.e., number of participants) within-subject measures of perceived task control for each of the five

error amplification methods. These repeated measures were used to statistically compare the effects of exercise with an error amplification method on changing the participants' perception of control over the task and thus, perception of task challenge (Table 2.2).

Figure 2.13 shows the overall trend in the level of perceived task control during practice with different error amplification conditions. The scores of self-reports to this measure showed a decrease as the EA methods changed from control to visual EA to visual plus force feedback, and from the lower gain to the higher. Post-hoc analysis of a within-subjects multivariate ANOVA showed that the means of dominance scores are significantly different between almost all pairs of EA conditions (at highest $p < 0.05$). The only exceptions are (1 pair): high-gain visual EA and low-gain visual plus force feedback EA methods.

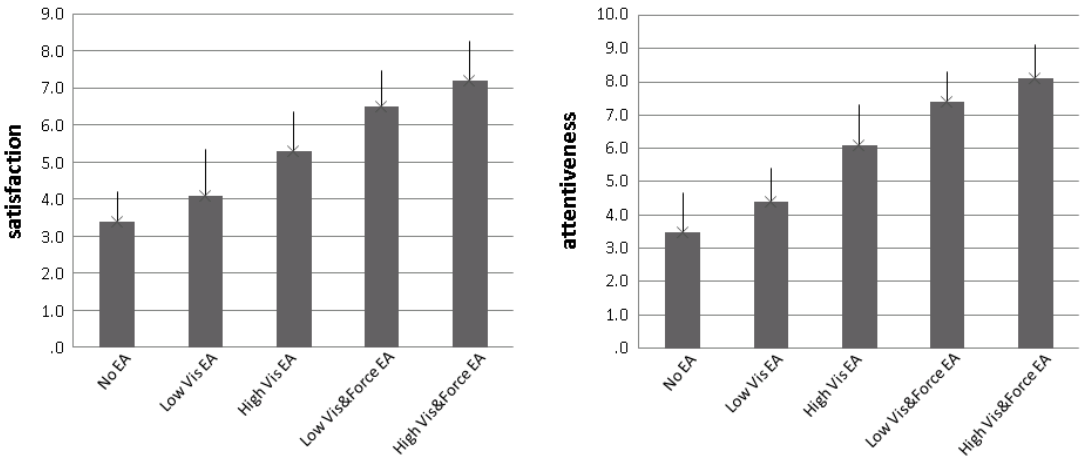


Figure 2.12: Overview of self-reports for satisfaction and attentiveness across error amplification types.

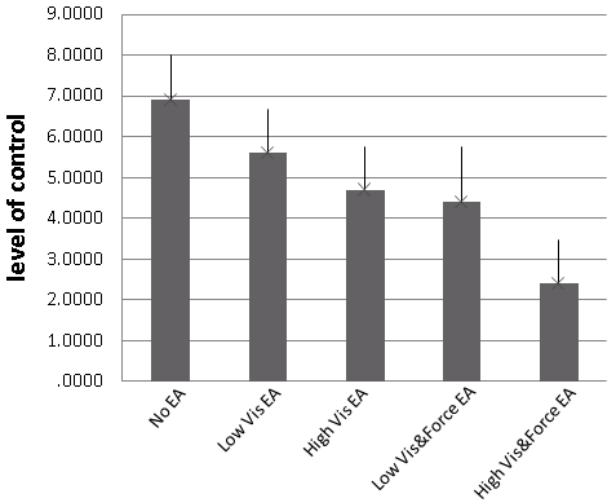


Figure 2.13: Overview of self-reports for perceived level of control across error amplification types

2.4 Discussion of Results and Limitations

The results presented in the previous section support both of the study hypotheses: 1) combining visual and force feedback error augmentation can improve motor adaptation (amount of improvement *a*) compared to control condition, 2) different error amplification methods are perceived as different levels of task difficulty.

The post-hoc analysis results presented in the previous section were corrected with the Bonferroni confidence interval adjustment method to deal with the problem of multiple comparisons and false positives (type I errors). The Bonferroni method is the simplest and most conservative method of correcting family-wise error rates. Mauchly's test of sphericity (a condition of data distribution that is required to conduct repeated measures ANOVA) was conducted to validate the repeated measures ANOVAs and all significant sphericity violations were corrected using the Greenhouse-Geisser method. Violation of sphericity refers to the case in which the variances of the differences between combinations of the ANOVA conditions are not equal [51, 52].

Comparing the means of adaptation speed and amount of improvement, low-gain visual-force feedback EA proved to be the best (Figure 2.10: higher amount of improvement *a* and faster speed of adaptation *b*), followed by high-gain visual and high-gain visual and force feedback EAs. Low-gain visual EA led to the worst adaptation pattern, suggesting that an amplification gain of 1.3 is not high enough to initiate learning, but it is high enough to make participants confused and decrease their performance. The ANOVA failed to find significant differences between the final performance "*c*" of participants after training with different EA methods, which is acceptable due to the fact that human motor function is not perfect and cannot fully follow ideal straight-line paths. In accordance with [19], the study presented in this chapter could not show significant differences between adaptation properties promoted by each EA, which the author believes could be reversed by a more careful tuning of the EA gains. The only exception was the high-gain visual and force feedback EA, which significantly improved the amount of improvement "*a*" in comparison to control and visual EA methods. This alone is not of high importance in terms of functional gain (the difference in final performance "*c*" is not significant for different EAs, different "*a*" values mean that different EAs cause different degrees of initial deviation). However, the difference in "*a*" values can be a contributing factor to increase satisfaction and attentiveness of the participants by inducing a feeling of improvement.

Several methodological differences between the Wei [18] and Celik [19] studies contributed to the fact that the ANOVA failed to show significant differences in the speed of motor adaptation measures "*b*" between EA methods. Wei and Celik used higher amplification gains (2 and 3.1) and compared error amplification using these gains with error offsetting (i.e., augmenting the cursor position by adding a constant value) in a between-subjects design. In these studies, adaptation was characterized using catch-trials (both distortion and EA are turned off: plain motion) placed in between exercise with challenge trials that ran for over 50 cycles. This study protocol design is well suited for studying only motor adaptation. In the study presented in this chapter, the low resolution of the robot's display (640 × 480) prevented using higher gains, such as 2 and 3.1, that were validated in other studies. The study reported here intentionally avoided using a between-subjects design with long and extended exercise blocks to simulate a condition close to an actual therapy session; a participant experiences different tasks in rather short periods of time.

The focus of this study was to propose error amplification methods as a way of adding "meaningful" difficulty "levels" to the robotic reaching task. "Meaningful" in this case means "leading to motor learning/adaptation". Results of this study support the idea that EA methods can be used as a meaningful tool for motor rehabilitation. A "level of difficulty" implies that the users are able to differentiate between the amount of challenge involved in doing a task at each of the levels.

Participants tended to be more satisfied with and more attentive during the visual plus force feedback EA methods, compared with the visual EA methods. The control condition had the lowest score for satisfaction and attentiveness. High satisfaction and attentiveness can be associated with high engagement in the task. Dominance followed a trend in the opposite direction, meaning that comparing the control condition with visual EA methods and also visual plus force feedback EA methods, the participants reported that the reaching task becomes more difficult and challenging to accomplish. It is arguable that a more difficult challenge in the reaching tasks compromises their repetitive nature and can thus lead to a higher satisfaction with the task. Similarly, completing a more challenging task requires a higher level of attentiveness.

This study investigated the changes in participants' affective states with only one question for each affect dimension (i.e., satisfaction, attentiveness, and dominance). Although the Self

Assessment Manikin questionnaire is a validated questionnaire, it is recommended to use multiple measures to evaluate human subjects' affective states.

2.5 Conclusion

The goal of this thesis is to apply the "desirable difficulties" methodology to robot-assisted reaching exercise and evaluate its efficiency in addressing the issue of engagement in stroke therapy. In a first step to this design process, this chapter presented trajectory error amplification as a way to promote motor adaptation, as well as to induce a perception of task difficulty. Results from this study provide strong statistical evidence, with at least 95% likelihood, that a participant will report that increasing the gain of error amplification and including additional means of amplifying error increases the difficulty of the reaching task.

The findings of this chapter support the idea that error amplification can be used as a way to introduce meaningful levels of difficulty into a robot-assisted reaching task. However, referring back to the main objective of this thesis to implement the "desirable difficulties" methodology in robot-assisted therapy, the following questions remain to be answered: Is it possible to predict a participant's desirable difficulty (desired level of error amplification)? How accurately can this be done? What is the best method to do so? And finally, what kind of information do we need to make such a prediction? These questions are investigated in an empirical study of machine learning methods in the next chapter.

3 Study II: An Empirical Study of Machine Learning Methods for Prediction of the User's Challenge Preference in A Robot-Assisted Reaching Task

Although robot-assisted rehabilitation regimens can be as effective, functionally, as conventional therapies, they still lack features to increase patients' engagement in the regimen. Providing rehabilitation tasks at a "desirable difficulty" is one of the ways to address this issue and increase the motivation of a patient to continue with a therapy program. In that case, one of the problems that needs to be addressed is to design a system that is capable of estimating the user's desirable difficulty, and ultimately, modifying the task based on this prediction.

This chapter presents the results of a study to compare the performance of three machine learning algorithms in predicting the direction of change to reach a user's desirable difficulty during a typical reaching motion rehabilitation task. Also, the usefulness of using participants' motor performance and physiological signals during the reaching task in prediction of their desirable difficulties was explored. Different levels of error amplification were used as different levels of task difficulty.

Section 3.2 gives an overview of the study goals and important factors that need to be considered in predicting desirable difficulties. Section 3.3 reviews the methodology of the study. Section 3.4 presents the results, followed by a discussion of the results in Section 3.5 and chapter conclusion in Section 3.6.

3.1 Introduction

As it was mentioned in the first chapter, recent neuroscience findings indicate that to stimulate brain plasticity and acquisition of new motor skills, intensive training and high dose of repetition are required [53]. The field of robot-aided stroke therapy has steadily progressed in the past two decades by incorporating this principle. Studies with different upper extremity rehabilitation robots such as MIT-Manus [11], Mirror Image Movement Enabler (MIME) [13], and GENTLE/s [15] have validated the positive effects of training with such devices on motor recovery and functional improvement. Additionally, in robotic therapy interventions, the rehabilitation process and the patient's progress can be quantified. This suggests the potential of these robotic devices to be utilized as a mainstream motor rehabilitation tool.

Research in both artificial intelligence and motor learning proposes an error-driven process that supports learning [44, 45] and can possibly maximize the effect of robot-assisted repetitive training. The previous chapter presented results showing that combining visual and force feedback error amplification with repetitive training improves the rate of motor adaptation in healthy individuals [54]. This can be generalized to the stroke population [18, 48].

Also, Chapter 1 discussed that while therapy is based on motor learning principles to increase a patient's motor function, prolonged engagement in the exercise is also a key factor in the success of a therapy regimen [22]. The mechanics of sustaining a user's engagement in a task has been theorized in the fields of game design and psychology. From a motor learning point of view, to avoid boredom or frustration, one needs to be kept at one's challenge point [29], or desirable difficulty, by meaningful manipulation of exercise difficulty. These desirable difficulties can be dependent on both task performance and a person's affective state. While quantifying task performance in robotic therapy regimens is relatively easy (Chapter 2), measuring affect is more challenging. Physiological signals processing to gauge affect has gained popularity in recent years. Chapter 1 provides a brief background of this emerging research field, and a more in-depth review follows.

Efficiency of several machine learning methods in predicting affect based on physiological signals is presented in [55]. In this study, Rani et al. used a human-robot interaction task to elicit different affective states. In a systematic comparison of the weaknesses and strengths of four machine learning algorithms, K-Nearest Neighbour, Bayesian Network, Regression Tree, and Support Vector Machine, the study showed that all the methods performed competitively.

Liu et al. [37, 56] have studied the relation between affective state and physiological responses. In separate studies, they proposed methods for prediction of a participant's affective state by using physiological responses and artificial intelligence. They showed that modifying tasks based on participant affect (e.g., anxiety and liking) leads to higher performance [37] and lower anxiety in an autistic population [56]. Wang et al. [57], using a performance-based assist-as-needed robotic training, have reported higher overall performance for participants training in an adaptive exercise program in comparison with a control group.

Novak et al. [58] used performance metrics and physiological measurements to design a biocooperative stroke rehabilitation system. In order to reduce dependency of robotic regimens to

human therapists to set up the tasks, Novak used different flavours of Discriminant Analysis to design a system that can predict a therapist's recommended changes for task difficulties. This work demonstrated greater performance of adaptive algorithms that can adjust their decision making process for each individual user compared to algorithms that work based on a generic model of the users. However, such systems will be more complex and computationally expensive.

Chapter 2 showed that doing a reaching exercise with different levels of error amplification leads to different levels of motor adaptation and affective states [54]. Moreover, participants reported different levels of perceived difficulty for each error amplification level. Building on these findings and using the methodologies of [55] and [58], the goal of this chapter is to investigate the potential of predicting the direction of change to reach a user's desirable difficulty. This chapter presents a comparison of three machine learning approaches with different degrees of complexity – Neural Networks, K-Nearest Neighbour, and Discriminant Analysis – in predicting participants' desirable difficulty in a robotic reaching task. These methods were applied to different input sets comprising of participants' motor performance and physiological signals. Results of this study can be used in future work to design a closed-loop control system that is capable of predicting the direction of change to reach the user's desirable difficulty to study the effects of exercising under desirable difficulty condition.

3.2 Methods

3.2.1 Research Ethics and Study Participants

This study was approved by the Clinical Ethics Research Board (CREB) of the University of British Columbia, and all participants provided written informed consent. Twenty-four healthy adult participants with an average age of 23.8 took part in this study. The male/female ratio of the participants was 12/12. In order to confirm participants were free of neurological impairment, a minimum score of 24 on the Folstein Mini-Mental Test was required. In addition, only participants with normal or corrected eyesight were recruited to make sure changes in participants' performance was not attributable to anything but changes in task difficulty (visual distortion and error amplification method).

3.2.2 Experimental Setup and Rehabilitation Task

Physiological signals were captured at 256 Hz using a ProCompInfiniti Physiology Suite (Thought Technology, Inc.). This study includes three physiological signals: skin temperature (Temp), skin conductance response (SCR), and respiration rate (Resp). Temperature was measured using a sensor strapped around the distal phalange of the ring finger of participant's dominant hand. SCR was recorded using two electrodes strapped around the distal phalanges of the index and middle fingers of the dominant hand. As in the previous study, participants controlled the robot with their non-dominant hand. The breathing rate sensor was placed on a strap around the participant's chest. Figure 3.1 shows a participant wearing the SCR sensor while doing the robotic rehabilitation task.

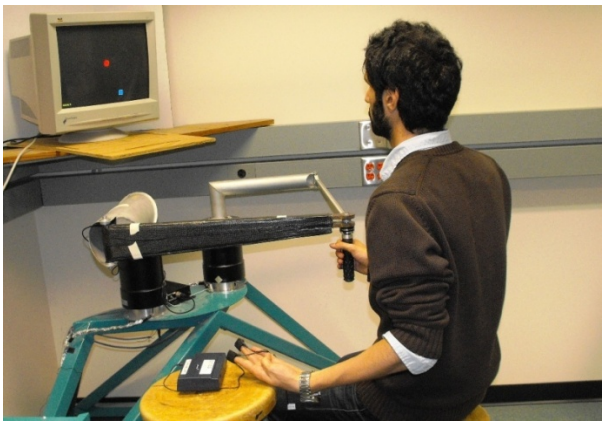


Figure 3.1: A participant holds robot's handle with his non-dominant hand while SCR data is being recorded from his left hand's fingertips

The custom-built five-bar robot that was used in the study presented in the second chapter was also used in this study. The robot was used to exercise reaching to targets in the horizontal plane as a motor learning task. Three targets were placed radially, with the same distance from the middle of the monitor and the other targets. When a target was presented, participants had to move the robot's handle with their non-dominant hand to place the moving dot over the target. Once the target was reached, they had to move the cursor back to the center of the screen. The term "cycle" is used for the consecutive reaching motions (i.e., trials) to the three targets. The order of the targets was predefined with a random number generator.

Participants were told the cursor represents the robot's end-effector position. The participants' hands were concealed with a cover in order to ensure that visual feedback was only provided via the monitor. Note that this cover is not shown in Figure 3.1.

To cause an initial deviation in healthy participants' reaching path and a decrease in their reaching performance, and thus to initiate motor learning and motor adaptation, a visual distortion was implemented during training blocks. This visual distortion was implemented as a ±30° rotation between the end-effector coordinates (i.e., actual hand position) and the monitor coordinates (i.e., visual target). Participants practiced with five different error amplification levels to learn reaching within this visual distortion.

Ordered from the easiest perceived difficulty by participants to the hardest (Chapter 2), the following five error amplification (EA) levels used were: control (no EA), low gain visual EA, high gain visual EA, low gain visual/force feedback EA, high gain visual/force feedback EA. More details about the robot-assisted reaching task, visual distortion and error amplification methods can be found in Chapter 2.

3.2.3 Data Collection Protocol

After providing consent to participate in the study, participants were introduced to the robotic manipulandum and the reaching task. Participants were then seated in front of the robot and physiological sensors were put on them. Each experiment, on average, took 80 minutes. Participants were given rest times (in addition to the existing rest periods) upon request. They were told that they could stop participating in the experiment at any point.

The experiment comprised of six exercise blocks (Figure 3.2, top). In the first block, participants practiced reaching in the virtual environment to become familiar with the task. This familiarization block comprised of 14 cycles in which both visual distortion and error amplification were turned off (i.e., plain motion). The familiarization block was followed by five training blocks (exercise blocks 2-6).

Each training block was divided into four sub-blocks (Figure 3.2, bottom): de-adaptation, rest period, exercise with challenge, and self-report of the direction of change to reach desirable difficulty. The de-adaptation sub-blocks comprised of 10 cycles of plain motion. This was to ensure training effects (i.e., adaptation to visual distortion) from previous training blocks did not carry over to the following challenge. In the rest period, participants were asked to remain seated and with the least possible motion for one minute. Physiological signals recorded in this period were used as a baseline in computing physiological measures. In an attempt to reduce possible distractions and physiological responses to these unwanted stimuli (i.e., noise in the

physiological data), the study room was kept quiet and the participants were asked to not initiate any conversations during these rest periods. Finally, the challenge sub-block comprised of reaching to visual targets within the visually distorted environment with one of the five EA levels (13 cycles). The order of appearance of the five EA levels throughout the training blocks (blocks 2 to 6) was random and changed for each participant.

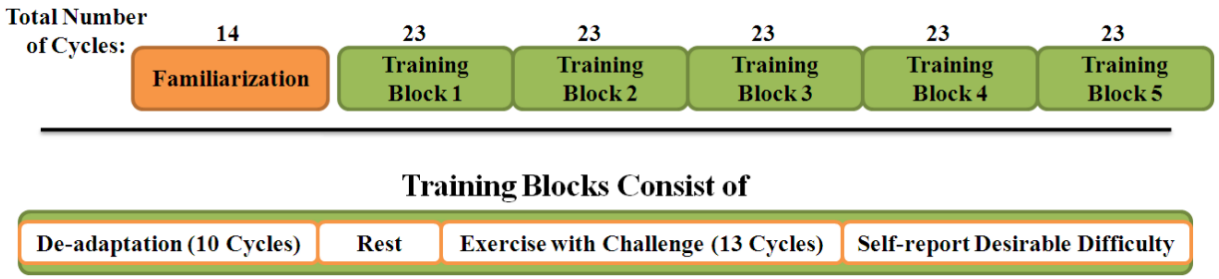


Figure 3.2: Top: Detailed break-out of the experiment protocol. Participants were given this information prior to the start of the experiment. Bottom: A break-out of a training block.

At the end of each of the training blocks (after finishing challenge sub-blocks), participants were asked to report whether they wanted their next trial to be easier (recorded as −1) or harder (+1), in other words, whether their "desirable difficulty" for the next training block was easier or harder than the challenge they had just practiced. To make the question clear, an example was given: "Assume that you want to play chess. If you play with a five-year-old, you will be winning all the time without feeling any challenge, and you will get bored. If you play with the world champion, you will be losing all the time without any hope for a win and that will make you frustrated. But, if you play with an average player the game will be engaging to you and you will enjoy the game. We call that desirable difficulty." The participants were not allowed to stay at the same level of difficulty, although it is likely that a participant finds a difficulty level as his/her *desirable* level of difficulty. Pilot studies showed that participants are inclined to stay at the same level of difficulty even if they are under- or over-challenged.

3.2.4 Motor Performance and Physiological Measures as Challenge Preference Prediction Attributes

Motor performance is measured by motor adaptation and reaching accuracy. As mentioned, a cycle consists of consecutive reaching trials to the three target points. For each cycle, the reaching accuracy measure is defined as the average of the maximum deviations of each of the three trials (Section 2.3.2). This maximum is calculated as the maximum deviation of the reaching trajectory from the line between the start point and the target point (i.e., the error vector

that is used for EA, section 2.3.1) in each trial. Figure 3.3 shows a plot of these maximum deviations over the course of exercise block number 2 (first training block) for participant 7. In this challenge sub-block, the participant trained with low gain visual/force feedback EA. As this dataset presents the first training block (no prior exposure to visual distortion), the de-adaptation period merely acts as more familiarization practice and there is no evidence of de-adaptation.

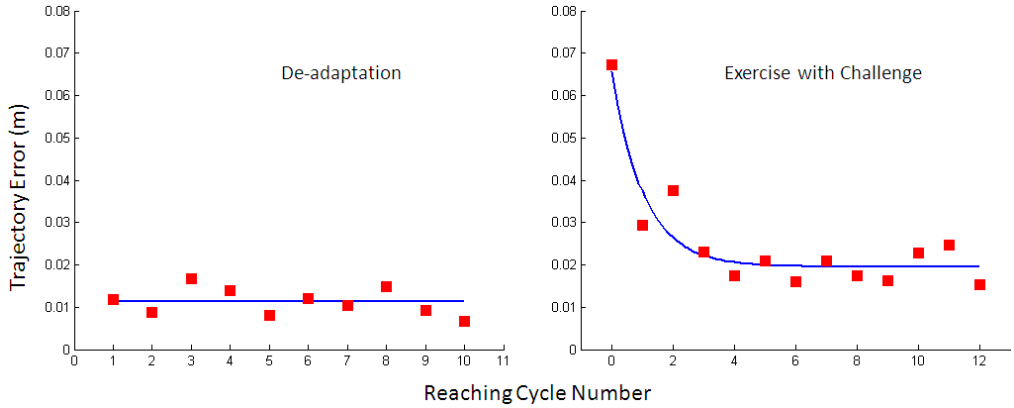


Figure 3.3: Left: following the familiarization block and during the first de-adaptation sub-block, the participant shows little variation in maximum trajectory deviation (about ±1mm). Right: during practice with challenge, as the participant adapts to the visual distortion with more practice, deviation decays.

As shown in Figure 3.3, right side, at the beginning of exposure to visual distortion, performance decreases (i.e., large initial deviation). With more practice, the participant adapts to the distortion and learns to perform reaching within the distorted environment with lower deviation from the straight line. As discussed in Chapter 2, adaptation can be modeled with an exponential decay function of this form:

$$y = ae^{-x/b} + c \tag{3.1}$$

In equation 3.1, y is the maximum trajectory deviation in a cycle, x is the cycle number, a is the total decrease in the maximum deviation (i.e., the difference between the maximum deviation in the first cycle and the last cycle), b is the time constant and represents the speed of adaptation, and finally c can be interpreted as the final maximum deviation after many cycles of practice (i.e., maximum deviation will eventually converge to c). For each challenge sub-block, a set of a, b, and c is calculated and used as 3 features of motor performance in predicting the direction of change to reach the participant's desirable difficulty.

During each rest period and challenge sub-block, three physiological signals were recorded: skin temperature (TEMP), respiration rate (RESP), and skin conductance rate (SCR) [32, 55, 58].

From these raw signals, 11 physiological features (those used in [58]) were extracted and used in predicting the direction of change to reach the participant's desirable difficulty. Figure 3.4 shows a plot of these raw signals over the course of exercise block number 2 (first training block) for participant 7. In this challenge sub-block, the participant trained with low gain visual/force feedback error amplification. Note the increase in the absolute value and variability of these signals progressing from the rest period to the exercise with challenge period.

From the SCR signal, 5 features were extracted: **normalized SCR**, defined as mean SCR in the challenge period divided by mean SCR in the rest period, **SCR peak value**, defined as the difference between minimum SCR and maximum SCR during the challenge period, **normalized derivative of SCR (dSCR)**, defined as mean dSCR in the challenge period divided by mean dSCR in the rest period, **SCR change**, defined as the difference between the mean SCR in rest period and mean SCR during the challenge period, and **dSCR variability**, defined as the standard deviation of dSCR in the challenge period.

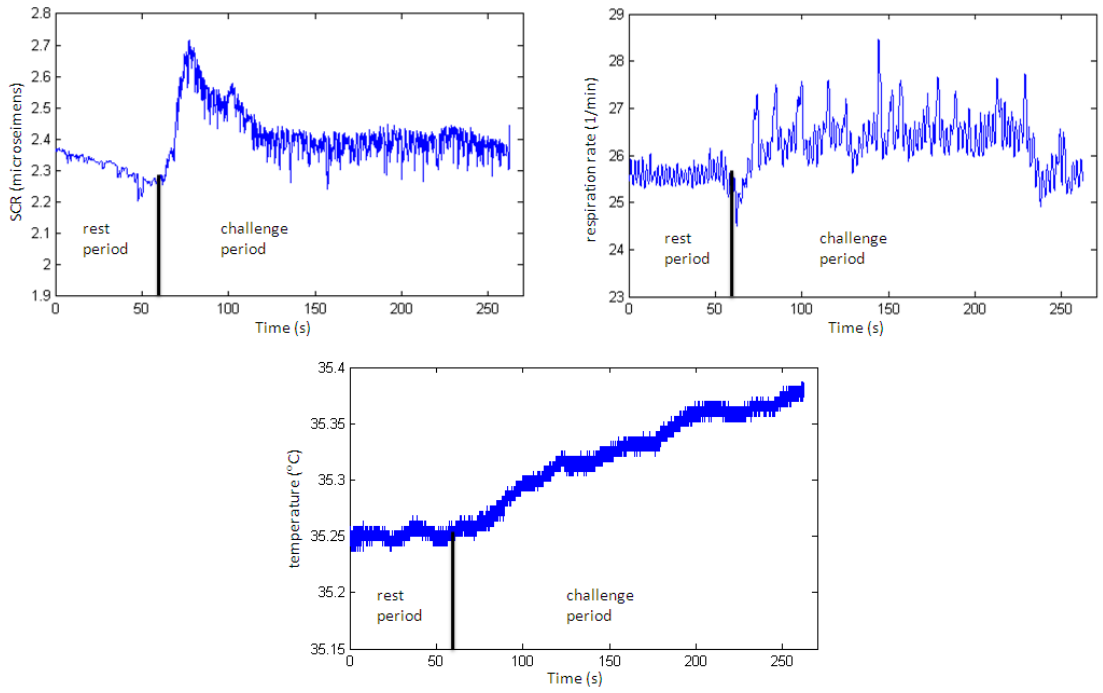


Figure 3.4: Raw physiological responses recorded from participant 7 during the first training block. Top-left: raw SCR signal. Top-right: raw respiration rate signal. Bottom: raw skin temperature signal.

From the respiration rate (Resp) signal, 3 features were extracted: **normalized respiration rate,** defined as mean Resp in the challenge period divided by mean Resp in the rest period, **respiration rate change,** defined as the subtraction of mean Resp during the rest period from

mean Resp during challenge period, and **respiration rate variability**, defined as the standard deviation of respiration rate in challenge period.

From the skin temperature (Temp) signal, 3 features were extracted: **normalized skin temperature,** defined as mean Temp in the challenge period divided by mean Temp in the rest period, **skin temperature change**, defined as the subtraction of mean Temp during the rest period from mean Temp during the challenge period, and **skin temperature variability**, defined as the standard deviation of skin temperature in the challenge period.

For each challenge sub-block completed by a participant, a set of 11 physiological features were extracted and used in predicting the participant's desirable difficulty.

3.2.5 Applied Machine Learning Algorithms

The goal of this study was to investigate the potential of predicting the reported direction of change to reach the desirable difficulty of the reaching task (easier or harder, −1 or +1) based on motor performance and physiological features. This was framed as a classification problem with the direction of difficulty change to achieve desirable difficulty (easier/harder) as the target function or output and motor performance and physiological features as the predictor variables or inputs. The relationship between the input and output datasets is non-linear and the high dimensionality of the input set increases the probability of having irrelevant variables in the input data. Then the two questions to be answered are: 1) What is the best machine learning algorithm to model the non-linear relationship between the input dataset and the target outcome? 2) Which of the performance and physiology metrics are better predictors of the direction of change to reach the user's desirable difficulty?

To answer these two questions, this study compared the accuracy of three machine learning algorithms in this classification problem by using four sets of predictor variables. The machine learning algorithms are k-Nearest Neighbour, Neural Network, and Discriminant Analysis, each having a different working principle and level of complexity associated with them.

The four sets of predictor variables are the motor performance features (3 attributes), physiological features (11 attributes), a hybrid of performance and physiological features (14 attributes in a one layer combination of the two feature sets), and a fuzzy combination (prediction based on the estimation confidence of each of the two feature sets). Comparing the

results of prediction based on each of these predictor variable sets informs us about the usefulness of including or excluding each of the feature sets from the prediction process.

The following is an overview of the machine learning algorithms and the fuzzy combination that are used in this study.

K-Nearest Neighbour (k-NN) Classification Method:

K-Nearest Neighbour (k-NN) is a non-parametric machine learning algorithm that requires storing of the entire training set. To classify a new instance, k (usually an odd number) nearest training samples to the new instance are considered to calculate a similarity score for the new instance. This similarity score is then used in assigning the new instance to one of the two outcome classes. k-NN is an improved variation of a lookup table and is sensitive to noisy and irrelevant data.

In this work, each new instance was assigned to the class with the highest sum of similarity scores (similarity score summing method):

$$C(X) = \text{argmax}_m \sum_{X_j \in \text{k-NN}} \text{Sim}(X, X_j) y(X_j, c_m) \tag{3.2}$$

In Equation 3.2, X is a new instance, m is the number of classes (here $m = 2$: easier or harder next trial), and X_j is one of the k neighbours in the training set. $y(X_j, c_m)$ is either zero or one and indicates if X_j belongs to class m (i.e., c_m) and Sim (X, X_j) is a measure of similarity calculated as the inverse of Euclidean distance between the two instances. The new instance is assigned to class m that has the highest sum of similarity scores with a probability p. p is calculated by dividing the sum of similarity scores for class m by the total of similarity scores for all classes. Different k values were tried to find a value that resulted in the highest prediction accuracy [59].

Neural Networks:

Neural Networks are a combination of weighted linear transformations that map the input space to the output space. A Neural Network (NN) is a parametric algorithm that does not require storing of the training data set.

Inspired by the structure and working principles of the human brain, a NN is a collection of units called neurons or nodes that are connected together by input and output links. A link propagates activation value a_i from unit i in a downstream node layer to unit j in a upstream node layer. Each

link has a numeric bias weight $w_{i,j}$. Unit j computes a weighted sum of its inputs (input j equals sum of all $w_{i,j} \times a_i$) and then applies an activation function g to its weighted sum of inputs to derive the output of the node. g is usually a logistic function (sigmoid perceptron).

During a learning phase, NN changes the values of its links' numeric bias weights $w_{i,j}$ to fit the relationship between the inputs and outputs. This learning of weights can be stated as an optimization problem. A method called back propagation of errors provides a mathematical solution to this problem.

Neural Networks is an active research area in the field of artificial intelligence. It is recommended to use Neural Networks with three layers of nodes [60]. The number of nodes in the first layer (input layer) is defined by the number of input variables. Nodes in this layer do not have an input link and each of them is set to output one of the input values. Additionally, each node in this layer is linked to all the nodes in the second layer. The second layer is called hidden layer and has at least the same number of nodes as the first layer [60]. Each of the sigmoid perceptron nodes in this layer is linked to all the nodes in the third layer. The third layer or the output layer also uses sigmoid perceptron nodes. The output value of node k in this layer can be used as a classification probability, showing the chance that the input set used belongs to class k. The highest probability decides the classification outcome.

Discriminant Analysis:

Discriminant Analysis (DA) is a parametric statistical method that uses a linear combination of attributes to calculate the probability of a data point belonging to a class. In DA the dependent variable (output) is categorical and the independent variables (input attributes) are continuous numerical quantities. Note that ANOVA and regression analysis are statistical methods that are used to express a *numeric* dependent variable as a linear combination of input attributes. Logistic regression is more similar to DA (both classifiers of categorical variables) and is preferable in applications where it is not reasonable to assume that the independent variables are normally distributed.

The discriminant function to classify data points into classes 1 and 2 can be formulated as [58, 59]:

$$D(X) = b + w^T . X \tag{3.3}$$

$$b = -w^T . \frac{1}{2} . (\mu_1 + \mu_2) \quad (3.4)$$

$$w = (S_1 + S_2)^{-1} . (\mu_2 - \mu_1) \quad (3.5)$$

$$C(X) = 1 \text{ if } D(X) < 0 \text{ or } 2 \text{ if } D(X) \geq 0 \quad (3.6)$$

In these equations, X is a new instance, $D(X)$ is a Discriminant function, μ_k is the matrix of the mean values of the input features for class k, S_k is the covariance matrix for class k, and $C(X)$ is the class to which X is assigned.

Fuzzy Combination of Performance-based and Physiological-based Predictions:

According to the theory of flow and the challenge point framework discussed in the first chapter, a user's task performance will initiate the engagement and learning process, which in turn will shape affective state (performance defines level of learning and engagement). This suggests that a simple hybrid of the performance and physiological metrics [58] might not be the best way to combine the two metrics. In a novel approach and to implement the main points of the flow theory and the challenge point framework, this thesis proposes a fuzzy combination of the performance and physiological metrics.

In this fuzzy combination approach, for each of the three machine learning algorithms, two decision making models based on the performance metrics and the physiological metrics are built. As mentioned before, all of the machine learning algorithms that are used in this study use a probabilistic approach in splitting data points into classes. New instances were fed to the performance-based model at first and if the performance-based classification outcome had a confidence greater than 70%, the class (as decided by the performance-based model) was reported. If the confidence of the performance-based classification was lower than 70%, the new instance would be fed to the physiological-based model. If the physiological-based classification outcome had a confidence greater than 70%, the class (as decided by the physiological-based model) was reported. If the confidence of the physiological-based classification was lower than 70%, the new instance would be classified randomly. This random assignment happened only twice during the training and validation phases.

3.3 Results

3.3.1 Training Machine Learning Algorithms

From the human-subject study (24 participants) a dataset with a total of 107 instances was collected. Each instance comprises of 14 numeric values for the predicting variables (i.e., 3 motor performance metrics and 11 physiological features as the input variables) and one self-report for the direction of change to reach desirable difficulty (–1 or +1 for an easier or harder next trial) as the outcome variable. The self-reports for an easier or harder next trial were split 45% to 55%.

To implement a k-fold cross-validation to compare the performance of different algorithms and input sets, the collected dataset was divided into 6 folds (i.e., $k = 6$). The recommended number of folds is 5-10 [59]. $k = 6$ was chosen to avoid a long training time. Four prediction models (one for each prediction variable set) were trained for each of the three machine learning algorithms using the data in each selection of 5 out of 6 folds ($4 \times 3 \times 6 = 72$ prediction models). The fold that was not used in training of the algorithms was used for validation of the trained models and comparison of their prediction accuracy.

To train the k-Nearest Neighbour models, $k = 7$ neighbours with Euclidean distances and equal distance weighting was used. After experimenting with different flavours of Discriminant Analysis methods, Diagonal Quadratic Discriminant Analysis yielded the best prediction accuracy. Finally, Neural Networks with one hidden layer were used. The number of nodes in the hidden layer that led to the best results (i.e., best training and validation accuracies resulted from a pre-test using hold-out cross-validation) is 6 nodes for performance-based model, 11 nodes for physiological-based model, and 14 nodes for the hybrid model.

After training each of the 12 prediction models for each selection of 5 out of 6 folds, the same instances of prediction variables that were used for training were fed to the algorithm. The number of correct predictions calculated from this process was used in defining the algorithm's ability to learn the training set. This was defined as a percentage and calculated as the ratio of the number of correct predictions divided by the total number of instances in the 5 folds. The results are presented in Table 3.1.

Table 3.1: Results of all prediction models' ability to learn the training sets (average accuracy ± standard deviation).

Machine Learning Algorithm	Prediction Variable Set			
	Performance features	***Physiological features***	***Hybrid of all features***	***Fuzzy combination***
K-Nearest Neighbour	79.0±1.1%	67.0±1.5%	75.2±1.3%	80.1±1.3%
Neural Network	80.1±1.0%	66.3±1.4%	77.2±1.5%	89.3±1.1%
Discriminant Analysis	76.8±1.4%	65.8±1.8%	75.6±1.5%	84.1±1.7%

3.3.2 Validation of the Trained Models

To validate the trained models and investigate their ability to generalize, in each of the six folds of training the instances in the fold that was not used in training of the algorithms were fed into each of the 12 models and their prediction accuracies were calculated (k-fold cross-validation). Results of this process are presented in Table 3.2.

To provide a common ground in comparing prediction accuracies, a random number generator as the baseline predictor model was used. Twenty sets of 107 random (−1, +1) pairs were generated and used as the output of the predictor for desirable difficulties. On average, the accuracy of randomly guessing the direction of change to reach a participant's desirable difficulty was found to be 49±2.3%, close to the expected value of 50%.

Table 3.2: Results of all prediction models' accuracy in predicting new validation instances (average accuracy ± standard deviation).

Machine Learning Algorithm	Prediction Variable Set			
	Performance features	***Physiological features***	***Hybrid of all features***	***Fuzzy combination***
K-Nearest Neighbour	70.2±2.5%	54.5±3.7%	67.5±2.3%	72.8±3.5%
Neural Network	72.8±2.6%	61.6±2.7%	71.0±2.0%	85.3±3.1%
Discriminant Analysis	71.8±2.9%	59.2±3.8%	67.5±1.63%	78.5±2.3%
Random Number Generator	49±2.3%[a]			

[a]Average of 20 trials using all 107 instances

3.4 Discussion

K-Nearest Neighbour (k-NN) is a simple non-parametric classification method that searches for similarities between the input data and the existing training data. However, this method showed the lowest ability to learn the training set using three of the variable sets (i.e., performance and

hybrid feature sets and fuzzy combination). On the other hand, it showed comparable results when it was tested with the validation dataset, especially using performance metrics or hybrid of all metrics. Neural Network (NN) is a parametric machine learning method. It showed the highest rate in learning the training dataset (except for the condition of physiological features as input), which resulted in better generalization of the data and thus better performance in the validation process. Discriminant Analysis (DA) showed slightly lower performance compared to NN (Figure 3.5).

Row by row comparison of the average accuracy values given in the Table 3.1 and 3.2 can be used to compare the performance of the three machine learning algorithms in training and validation phases using different input sets. Although the three methods showed different levels of ability to learn the dataset, they demonstrated competitively close accuracies in predicting desired difficulties based on new data. NN, followed by DA and k-NN, had the best average accuracies in learning the training dataset and predicting the direction of change to achieve participants' desirable difficulties in the validation phase.

To evaluate the trained models and assess their ability to generalize to new input data, the k-fold cross-validation method was used. Leave-one-out cross-validation (LOOCV) is more common in the evaluation of machine learning methods used for affect recognition or prediction of a user's preference [55, 58]. However, given the dataset size, use of LOOCV would be more computationally expensive compared to the use of the k-fold cross-validation method (training 107 models vs. 6 models for each of the 12 conditions). Additionally, the k-fold cross-validation method gives a more objective estimate of accuracy and generalizability by providing both average and standard deviation of prediction accuracy.

To assess the usefulness of the four different prediction variable sets, results presented in Table 3.2 were compared column by column. Prediction based on motor performance features has a high accuracy rate (highest 72.8%) among all three machine learning methods, and it is followed by prediction based on a hybrid of all the features (highest 71%) and then physiological features alone (highest 61.6%). The fuzzy combination has the highest accuracy rate in comparison to the other input sets (highest 85.3%). This overall trend for the usefulness of different input sets in improving prediction accuracy is also demonstrated in Figure 3.5.

The machine learning algorithms used in this study are the most popular methods of predicting a user's preference or affective state in the literature [32, 55, 58]. However, there has not been a study to systematically compare these methods in predicting a user's preference. Note that the Neural Network and Discriminant Analysis methods can optimize their structure to give the least value to the attributes that do not contribute to the prediction of the target function (non-relevant attributes) and thus, exclude them from the prediction process. On the other hand, the k-Nearest Neighbour method performs poorly if part of the input data is noisy or non-relevant. Therefore, the lower accuracy of prediction based on physiological features alone can be attributed to the noisy nature of these features (compared to the motor performance features). Future work can include investigation of the usefulness of each of the 11 physiological features introduced in this study, in order to exclude the non-informative and non-relevant features in training the machine learning algorithms.

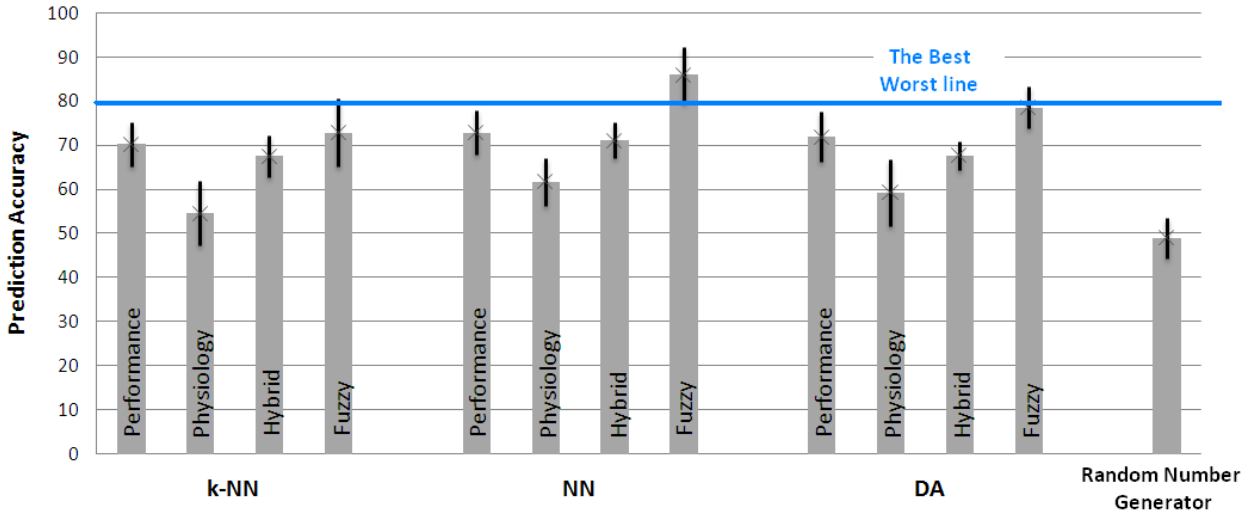


Figure 3.5: Comparison of all 12 combinations of input sets and machine learning algorithms. Error bars show two standard deviations (95% confidence intervals). The most conservative estimate of the best method (best worst line) is the combination of Neural Network and fuzzy combination of the features.

The noticeable superiority of prediction based on motor performance compared to physiological features (73% vs. 62%) may suggest the possibility of simplifying the robotic therapy system by removing the physiological sensors and avoiding the physiological signal analysis altogether. However, the results of the fuzzy combination of the two feature sets prove the usefulness of this type of physiological computing and the fact that physiological signals can provide supplementary information in predicting the direction of change to reach a user's desirable difficulty. This can be demonstrated by drawing the best worst line for prediction accuracy (using a 95% confidence interval) between all of the 12 combinations of input sets and machine

learning algorithms (blue horizontal line in Figure 3.5). This line passes through the lowest expected accuracy of prediction using Neural Network algorithm and fuzzy combination of input features. This line is above the performance ranges of all of the other three input sets and only crosses the performance ranges of prediction using DA and k-NN with fuzzy combination of input features. However, as the computational expenses of the three machine learning algorithms are the same, prediction using a Neural Network approach and the fuzzy combination of the input features can be the most appealing combination for further investigation (i.e., closed-loop studies to measure the effects of practicing under desirable difficulties). This is due to the fact that this combination has the highest average prediction accuracy.

Nevertheless, a more conservative test to choose the input set plus algorithm combination that leads to higher accuracies would be to define the significance of the differences between the 12 combinations by running ANOVA tests. This analysis was replaced by the best worst line analysis as the k-fold cross-validation with $k = 6$ does not produce enough sample points (i.e., 6 accuracies per condition; recommended minimum sample size for ANOVA tests $N = 10$ [51]). Performing a k-fold cross-validation with $k = 10$ would excessively increases the computation time.

Future work following this phase will include the exploration of more sophisticated machine learning methods such as Support Vector Machines (SVM). SVMs have three properties that make them a popular choice [59]: 1) the ability to construct a maximum margin separator (i.e., decision boundaries with the largest possible distances from training instances) that improves the generalizability to new instances, 2) using the "kernel trick", SVMs embed the data into a higher dimensional space and create a linear separating hyperplane. This means that SVMs are able to expand the hypothesis space, and 3) being resistant to overfitting. SVMs have been used in an affective computing and human-robot interaction study by Rani et al. [55], yielding an accuracy rate of 85% in classifying human-subjects' affective states.

3.5 Conclusion

To investigate the possibility of estimating the direction of change to reach a user's desirable difficulty during a rehabilitation exercise, this chapter presented the accuracy of three commonly used machine learning algorithms in a comparative study. Data from a human subject study with 24 healthy participants were used to train models that map participants' motor performance and physiological signals to their desirable difficulties. To elicit different responses (i.e., motor

performance and physiological changes) participants were asked to exercise with a robotic rehabilitation reaching task. From the most accurate to the least accurate, the three machine learning algorithms are Neural Network, Discriminant Analysis, and k-Nearest Neighbour. All three methods performed competitively and showed superior performance in comparison with random prediction. Additionally, results showed use of a fuzzy combination of motor performance and physiological attributes yields higher prediction accuracy compared to using only one of the two input sets. The best prediction accuracy was achieved by using a Neural Network model and fuzzy combination of the motor performance and physiological metrics.

The findings of this chapter support the idea that it is possible to predict the direction of change to reach a user's desirable difficulty for the next training task with a high accuracy. However, going back to the goal of this thesis to implement the "desirable difficulties" methodology in robot-assisted therapy, the following questions remain to be answered: Does practicing under desirable difficulty condition lead to a more positive training experience for the users? What are the measures to quantify this possible positive experience? These questions are investigated in a between-subjects study in the next chapter.

4 Study III: Evaluating the User's Experience while Exercising under Adaptive Desirable Difficulty Conditions

The results of the studies presented in the last two chapters show that error amplification can be used as a way to introduce levels of challenge into a robotic reaching task, and the direction toward a user's desired level of challenge can be estimated with 85% accuracy. Building on those findings, this chapter presents the results of a study to investigate the impacts of practicing under desirable difficulty conditions.

Section 4.1 reviews the purpose of this study, followed by study methods in Section 4.2. Section 4.3 presents the results of this study, and finally Section 4.4 discusses the results and concludes the chapter.

4.1 Introduction

4.1.1 Overview and Purpose of the Study

The first chapter presented a methodology to increase users' engagement in a motor learning task. To achieve higher task engagement, this methodology involved delivering physical exercise at the user's desirable difficulty level. As discussed in Chapter 1, implementing this method involves two main steps: introducing meaningful levels of difficulty into the physical exercise and designing a system that is capable of estimating the user's desirable difficulty (to deliver the exercise at the user's desired level of challenge).

To investigate the feasibility of implementing this methodology, this thesis has focused on the practice of reaching motions using a robotic manipulandum. Chapter 2 discussed the effects of practicing reaching motions with error amplification. The study presented in that chapter showed that different levels of error amplification lead to motor adaptation and that these levels were experienced by the study participants as different levels of task difficulty. Chapter 3 presented Study II, an empirical evaluation of different methods of predicting the direction of change to reach a user's desirable difficulty level. The results of that study showed that this direction of change for the next training block can be estimated by a pre-trained Neural Network algorithm up to 85% of the time. These two studies provide strong empirical evidence that the "desirable difficulties" methodology can be implemented in a robotic reaching task.

However, these studies did not explore the impacts of practicing the task under desirable difficulty conditions. Considering the larger goal of increasing task engagement during a motor learning task, this chapter investigates whether a user will have a more positive perception of the robot and the physical exercise when the robot delivers the reaching exercise at the user's desirable difficulty. This positive perception covers different affective states such as motivation and engagement, as well as the users' opinion about the task and robot.

In a between-subjects study, participants completed 5 training blocks of reaching motions. Participants in a control group received the training blocks in a predefined random order while participants in an experimental group received the training blocks based on predictions of the participant's desirable difficulty performed by a machine learning algorithm. To quantify the experience of the two groups, the SAM questionnaire (introduced in Chapter 2) and a post-experiment questionnaire were used. Participants' responses to these two questionnaires were studied to investigate if the experience of the robot-assisted training session as well as the users' perception of the robot and the task is different between the two groups. The following section provides more details about the use of these questionnaires.

4.1.2 Questionnaire Design

There are several standard questionnaires used in the two neighboring fields of Human-Robot Interaction (HRI) and Human-Computer Interaction (HCI) to gauge human experience and human perception of tasks and tools (i.e., interfaces). This study was specifically interested in three items of perception: 1) users' perceptions of the reaching task and the task load, 2) users' perceptions of the robot's responsiveness and usefulness as a trainer, and finally 3) users' perceptions of their experience of the entire training session. These human perception factors were measured using a post-experiment questionnaire and the SAM questionnaire.

The U.S. National Aeronautics and Space Administration (NASA) has developed a subjective workload assessment tool known as the NASA Task Load Index (NASA-TLX) [61, 62]. This tool has been used to evaluate workload in various human-machine systems in several different contexts. NASA-TLX is a multi-dimensional rating tool that measures six metrics using five-point Likert scale questions. These metrics are: mental demand, physical demand, temporal demand, own performance, effort, and frustration. To measure users' perceptions of the reaching task and the task load, five of the six subscales of the NASA-TLX were included in the post-experiment questionnaire. The temporal demand subscale was not included in this study as the

timing of the task was not a variable between the two participant groups (from the beginning of the reaching motion, a participant has only 0.5 seconds to complete the task; see Section 2.2.3 for details).

Users' perceptions of the robot were measured by using two components of the God-speed questionnaire [63], a standard HRI questionnaire. Five-point Likert scale questions were used in the post-experiment questionnaire to measure the robot's usefulness and responsiveness.

O'Brien has studied the process of user engagement in HCI [64] and has proposed a universal instrument for measuring user engagement called the User Engagement Scale [65]. To quantify the user experience over the entire training session, five metrics of the User Engagement Scale were used in the post-experiment questionnaire presented as five-point Likert scale questions: the task's interestingness (engagingness), the user's willingness to continue the training, the user's focused attention, motivation, and control and feedback.

In addition to the post-experiment questionnaire, the Self Assessment Manikin (SAM) questionnaire [50] was used to gauge the level of change in participants' satisfaction with and attentiveness during the task from the beginning of the experiment to the end of it.

4.2 Methods

4.2.1 Research Ethics and Study Participants

This study was approved by the Clinical Research Ethics Board (CREB) of the University of British Columbia, and all participants provided informed written consent. Twenty-three healthy adult participants with an average age of 24.3 and a male/female ratio of 11/12 took part in this study. In order to confirm participants were free of neurological impairment, a minimum score of 24 in Folstein Mini-Mental Test was required. Moreover, only participants with normal or corrected eyesight were recruited to make sure changes in participants' performance were not attributable to anything but changes in the presence of visual distortions and the error amplification methods.

The study participants were put into a control group and an experimental group: 13 participants in the control group and 10 participants in the experimental group. The male/female ratio in the control group was 6/7 and in the experimental group was 4/6. The age averages of the two groups were 24.2 years for the control group and 24.3 years for the experimental group. For the

experimental group, error amplification level of the reaching exercise varied based on predictions of a trained machine learning algorithm. The level of error amplification was varied randomly for the control group.

4.2.2 Experimental Setup and Motor Adaptation Task

This study employed the same experimental setup as for the study presented in Chapter 3. The two main devices used in this study were the robotic manipulandum to deliver the reaching task with different levels of error amplification and a suite of physiological signals sensors to collect participant physiological responses.

Physiological signals were captured at 256 Hz using a ProCompInfiniti Physiology Suite (Thought Technology, Inc.). Three physiological signals were included in this study: skin temperature, skin conductance response and respiration rate. The positions of physiological signals sensor straps on participants' bodies were similar to those of the study presented in Chapter 3. The physiological responses were used in determining participants' levels of desirable difficulty.

The custom-built five-bar robot used in the studies presented in the previous two chapters was also used in this study. The robot was used to exercise reaching to targets in the horizontal plane as a motor learning task. Three targets were placed radially (equally spaced), with the same distance from the middle of the monitor. When a target was presented, participants had to move the robot's handle with their non-dominant hand to place the moving dot over the target. Once the target was reached, they had to move the cursor back to the center of the screen. The order of target presentation was predefined with a random number generator.

The participants' non-dominant hand was concealed with a cover in order to ensure visual feedback was only provided via the monitor. Participants were told the cursor represents the robot's end-effector position. To cause an initial deviation in healthy participants' reaching path and to initiate motor adaptation, a 30° rotation was used as visual distortion. Ordered from the easiest perceived difficulty to the hardest (Chapter 2), the five error amplification (EA) levels used were: control (no EA), low gain visual EA, high gain visual EA, low gain visual/force feedback EA, high gain visual/force feedback EA.

More details about the physiological signals sensors, the robot-assisted reaching task, visual distortion, and error amplification methods can be found in Chapters 2 and 3.

4.2.3 Adjusting Training Block Challenges Based on Prediction of Participants' Preference

The study presented in Chapter 3 investigated methods of desirable difficulty prediction and the most informative inputs for this process in an open-loop validation process. The results showed that using Neural Networks algorithm and a fuzzy combination of the motor performance and physiological metrics (see Section 3.2.4 for definitions of these metrics) leads to the highest prediction accuracy.

In the open-loop validation study presented in the previous chapter, a k-fold cross-validation method with $k = 6$ was implemented. In this study, all 107 data points that were collected in the study presented in Chapter 3 were used in a fuzzy combination to train a new Neural Network. After training this new Neural Network (NN) model, the same 107 instance set that was used for training was fed back to the NN model. The model's prediction of the direction of change to achieve desirable difficulty for 96 of the data points matched the known direction of desirable difficulty. The number of correct predictions can be used to define the algorithm's ability to learn the training set, a measure for goodness of training. This is defined as a percentage and calculated as the ratio of the number of correct predictions (i.e., 96) divided by total number of predictions (i.e., 107), which in this case equals 89%.

This new Neural Network model was used for predicting the direction of change to achieve desirable difficulties for the experimental group in this study.

4.2.4 Data Collection Protocol

After providing consent to participate in the study, participants were introduced to the robotic manipulandum and the reaching task. Participants were then seated in front of the robot, and the physiological signal sensors were put on them. Each experiment, on average, took 80 minutes. Participants were given rest times at any time upon request (in addition to the scheduled rest periods in the exercise blocks). They were told that they could stop participating in the experiment at any point.

The data collection protocol used for this study was similar to that of the study presented in Chapter 3. The experiment comprised of six exercise blocks (Figure 4.1, top). In the first block, participants practiced reaching in the virtual environment to become familiar with the task. This familiarization block comprised of 14 cycles in which both visual distortion and error

amplification were turned off (i.e., plain motion). The familiarization block was followed by five training blocks (exercise blocks 2 to 6).

Each training block was divided into four sub-blocks (Figure 4.1, bottom): de-adaptation, rest period, exercise with challenge, and self-report of desirable direction of change in difficulty. Similar to the previous studies, the goal of the de-adaptation sub-blocks was to remove training effects of previous training blocks and each de-adaptation sub-block comprised of 10 cycles of plain motion. In the rest period, participants were asked to remain seated and with the least possible motion for one minute. Physiological signals recorded in this period were used as a baseline in computing physiological measures. In an attempt to reduce possible distractions and physiological responses to any unwanted stimulus, the study room was kept quiet.

The "exercise with challenge" sub-block comprised of reaching to visual targets within the visually distorted environment with one of the five EA levels (13 cycles). For the control group, the order of appearance of the five EA levels throughout the training blocks was random and changed for each participant. However, for the experimental group, only the EA level of the *first* training block was chosen randomly; the level of EA for the rest of the training blocks was changed based on the predictions of a trained machine learning algorithm. The Neural Network algorithm with fuzzy combination of motor performance and physiological metrics validated in the previous chapter was employed for prediction of desirable direction of change in difficulty in this study. At the end of each training block, the machine learning algorithm was used to predict whether the user's desirable difficulty for the next training block should be easier or harder than the challenge of the just-completed training block. Based on this prediction, the appropriate EA level was implemented for the next training block.

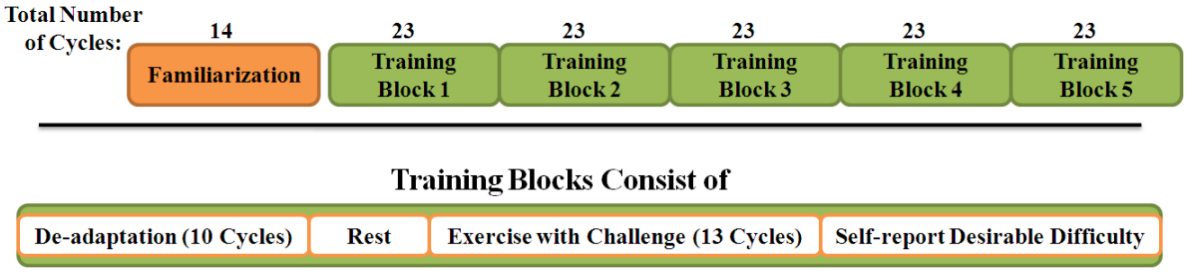


Figure 4.1: Top: Detailed break-out of the experiment protocol. Participants were given this information prior to the start of the experiment. Bottom: A break-out of a training block.

At the end of each of the training blocks (after finishing the challenge sub-blocks), all participants were asked to report whether they wanted their next trial to be easier (recorded as

−1) or harder (+1), in other words, whether their "desirable difficulty" for the next training block was easier or harder than the challenge they just practiced. To make the question clear, the same example used in the study presented in Chapter 3 was given. The answers of the experimental group to this question were compared with the output of the desirable change in difficulty prediction algorithm for closed-loop validation of the machine learning method.

4.3 Results

4.3.1 Demographics

Twenty-three participants were involved in this study: 13 participants in the control group and 10 participants in the experimental group. The male/female ratio in the control group was 6/7 and in the experimental group was 4/6. A t-test failed to show any difference between the almost identical age averages of the two groups: 24.2 years for the control group and 24.3 years for the experimental group.

Participants in both groups were asked to report their level of familiarity with robots on a scale of 1 (not familiar) to 5 (very familiar). The control group on average reported a score of 1.92 for familiarity with robots. This score was 1.80 for the experimental group. A t-test failed to show a significant difference between the averages of the reported values for familiarity with robots between the two groups. This suggests that the effects of robotic task novelty is same for both groups and possible differences in the reported scores for human perception of the task and robot cannot be attributed to participants' level of familiarity with robots.

4.3.2 Closed-loop Validation of the Machine Learning Algorithm for Prediction of Desirable Difficulties

Using the Neural Networks algorithm with a fuzzy combination of the motor performance and physiological metrics yielded the highest prediction accuracy in the open-loop hold-out cross-validation presented in the previous chapter. As mentioned earlier, a new Neural Network (NN) model was trained for this closed-loop study (training accuracy of 89%). A closed-loop validation of the new model was performed to demonstrate that this method of prediction can be used as an online decision making feedback loop for task adaptation.

Each of the 10 participants in the experimental group completed 5 training blocks and after each block reported whether their desirable difficulty for the next block was easier or harder than the

challenge they just completed. The self-reports of the experimental group for an easier or harder next trial were split 44% to 56% (45% to 55% for the control group).

This produced 50 new data points: a pair of motor performance and physiological metrics plus self-report of challenge preference. The new NN model was used to predict the participants' preferred change of difficulty for on-line adjusting of the next training block. These predicted challenge preferences were compared against the reported challenge preferences to perform a closed-loop validation of the trained NN model.

Overall, the new NN model's prediction was correct for 39 of the 50 new data points (78%). Considering the participants individually, the accuracy ranged from a minimum of 3 out of 5 correct predictions for participants #16 and #18, to 5 out 5 correct predictions for participant #17 (Figure 4.2).

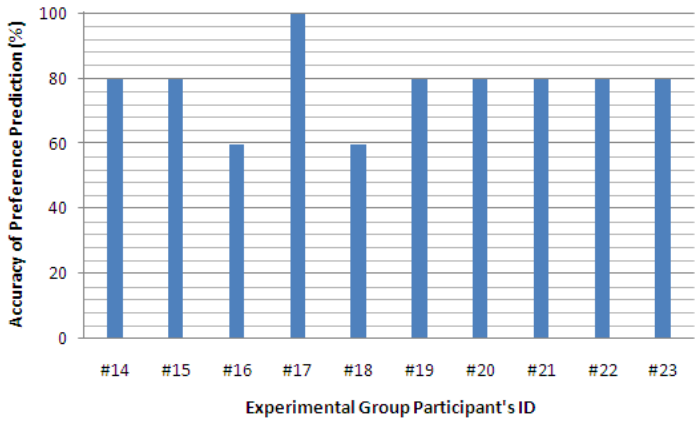


Figure 4.2: Accuracy of the machine learning algorithm in predicting desirable difficulties for the participants in the experimental group.

4.3.3 Participants' Perceptions of the Reaching Task and the Task Load

The two groups' perception of the reaching task and the task load was collected using five subscales of the NASA Task Index: mental demand, physical demand, task performance, perceived effort, and frustration. The mean and standard deviation of these five measures are reported in Table 4.1 and presented graphically in Figure 4.3.

Summarized in Table 4.2 are the results of pair-wise comparisons of the five task perception measures between the control and experimental groups. The pair-wise comparison was done using a two-tailed *t*-test and assuming unequal variances (a condition of the data in this study). Performing *t*-test with this assumption involves adjusting degrees of freedom *df* to a lower value, which increases the critical *t*-value for $p < 0.05$ and makes the test more conservative [51, 52].

Table 4.1: The means and standard deviations, in parentheses, of the participants' perception of the task and task load measures for each of the participant groups. Measures showing trending or significant differences between the two groups are indicated with the following suffixes: ^{t} $p < 0.10$, * $p < 0.05$.

Measure	Control Group	Experimental Group
Mental Demand	2.84 (1.21)	2.80 (1.13)
Physical Demand	2.69 (1.18)	2.40 (1.34)
Task Performance *	2.54 (0.96)	3.40 (0.97)
Perceived Effort *	3.85 (0.89)	3.00 (0.94)
Frustration ^{t}	3.15 (1.28)	2.30 (0.95)

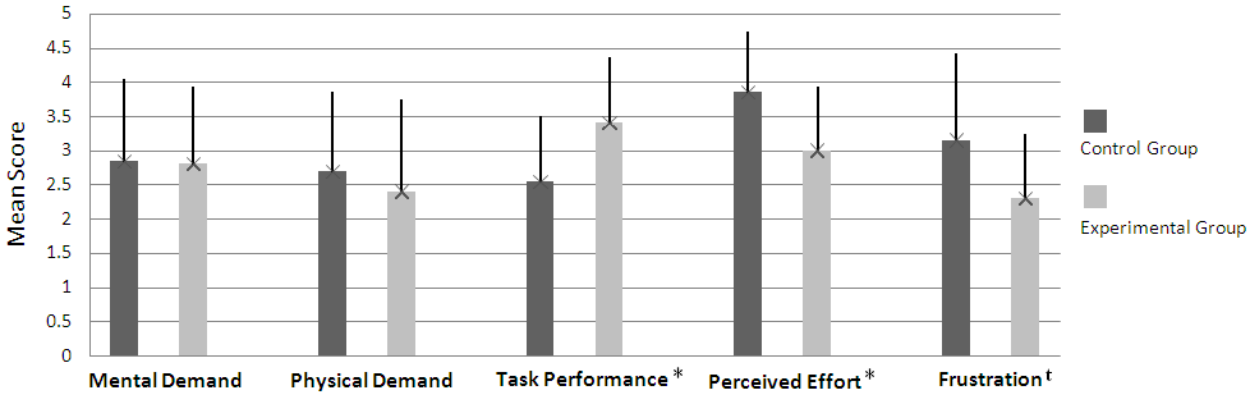


Figure 4.3: Overview of the participants' perception of the task and task load scores collected from five-point Likert scale questions. Measures showing trending or significant differences between the two groups are indicated with the following suffixes: ^{t} $p < 0.10$, * $p < 0.05$.

Table 4.2: Two-tailed *t*-test results for pair-wise comparison of the participants' perception of the task and task load measures between the two participant groups. Measures showing trending or significant differences between the two groups are indicated with the following suffixes: ^{t} $p < 0.10$, * $p < 0.05$.

Measure	*t*-value (*df*)	*p*-value
Mental Demand	0.09 (20)	0.92
Physical Demand	0.54 (18)	0.59
Task Performance *	2.12 (20)	0.04
Perceived Effort *	2.17 (19)	0.04
Frustration ^{t}	1.83 (21)	0.08

The results of pair-wise comparison of the reported scores using two-tailed *t*-test failed to show any significant difference in the reported scores for perceived mental and physical demands of the task between the two groups. Additionally, the results also showed that the experimental group reported a higher score of task performance and a lower score for perceived effort compared to the control group. Both of these differences are statistically significant. Although this analysis demonstrated a trend in the score for frustration (lower for experimental group), the *t*-test failed to prove the significance of the trend seen in the data ($p = 0.08$).

4.3.4 Description of the Robot

The Godspeed and the User Engagement Scale questionnaires were used in this study to collect metrics that quantify participants' perceptions of the robot and the training session (Section 4.1.2). Both questionnaires were designed to collect multiple metrics of the same measure in order to increase the trustworthiness of the data. In this study, the metrics of a same measure were aggregated and their mean value was used as the score for that measure. Thus, it was required to ensure that the collected data for each measure is internally reliable. This was done by calculating Cronbach's alpha value, an estimate that shows the consistency of the reported values for different metrics that quantify one measure (Table 4.3).

Table 4.3: Internal reliability of the measures used to quantify participants' descriptions of the robot and the training session.

	Measures	Cronbach's alpha	Metrics
Description of the Robot	Usefulness	0.709	Independent (reverse), Helpful
	Dominance	0.712	Assertive, Competitive, Dominant, Forceful
Description of the Experience of the Entire Interaction	Engagingness	0.776	Boring (reverse), Enjoyable, Engaging
	Willingness to Continue	N/A	If time was not a constraint, how many more training blocks would you do?
	Focused Attention	0.522	Lost myself in the experience, I lost track of time
	Motivation	0.843	The experience was fun, I felt discouraged during the interaction (reverse), I felt frustrated during the interaction (reverse)
	Feedback	0.554	I felt involved in the experience, The interaction did not work as I planned (reverse), I felt in control of the experience

Table 4.3 summarizes the metrics that were used to quantify different measures and the value of Cronbach's alpha for each measure. Measures with Cronbach's alpha value below 0.7 are not considered internally reliable and were excluded from further analysis (i.e., *t*-test). Both of the measures for quantifying participants' descriptions of the robot – robot's usefulness and robot's dominance over the participant (reverse of responsiveness) – were internally reliable. The mean and standard deviation of these two measures are reported in Table 4.4 and presented graphically in Figure 4.4.

Table 4.4: The means and standard deviations, in parentheses, for the reported scores of the robot's usefulness and dominance for each participant group. Measures showing significant differences between the two groups are indicated with the following suffix: ** $p < 0.01$.

Measure	Control Group	Experimental Group
Usefulness**	2.31 (0.85)	3.40 (0.94)
Dominance	3.31 (0.56)	2.92 (0.80)

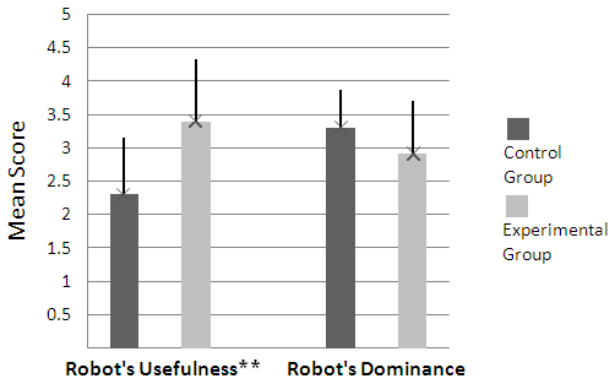


Figure 4.4: The reported scores of the robot's usefulness and dominance for each participant group collected from five-point Likert scale questions. Measures showing significant differences between the two groups are indicated with the following suffix: ** $p < 0.01$.

Summarized in Table 4.5 are the results of pair-wise comparison of the two descriptions of the robot measures between the control and experimental groups. The pair-wise comparison of the two participant groups was done using a two-tailed *t*-test and assuming unequal variances. Although the difference in the scores of the robot's usefulness is statistically significant, the difference in the dominance scores is not.

Table 4.5: Two-tailed *t*-test results for pair-wise comparison of the reported scores of the robot's usefulness and dominance for each participant group. Measures showing significant differences between the two groups are indicated with the following suffix: ** $p < 0.01$.

Measure	*t*-value (*df*)	*p*-value
Usefulness**	2.88 (19)	$p < 0.01$
Dominance	1.29 (15)	$p = 0.22$

4.3.5 Description of the Interaction

As mentioned above, Table 4.3 summarizes the five measures that were used to quantify participants' experience during the five training blocks. Only three of the measures (engagingness of the experience, willingness to continue the exercise, motivation) were internally reliable, while the other two (focused attention and available feedback) had a Cronbach's alpha value lower than 0.7 and were excluded from the subsequent pair-wise

comparison of the two groups. The mean and standard deviation of these five measures are reported in Table 4.6 and presented graphically in Figure 4.5.

Table 4.6: The means and standard deviations, in parentheses, for the measures quantifying participants' experience over the duration of the five training blocks. Measures of focused attention and feedback were not internally reliable and t-tests were not performed on them. Measures showing significant differences between the two groups are indicated with the following suffix: * $p < 0.05$, *** $p < 0.001$.

Measure	Control Group	Experimental Group
Engagingness *	3.13 (0.99)	3.83 (0.57)
Willingness to Continue ***	1.23 (1.48)	4.60 (1.83)
Focused Attention	2.62 (1.02)	2.80 (0.97)
Motivation *	3.23 (1.13)	4.03 (0.61)
Feedback	3.23 (0.90)	3.67 (0.68)

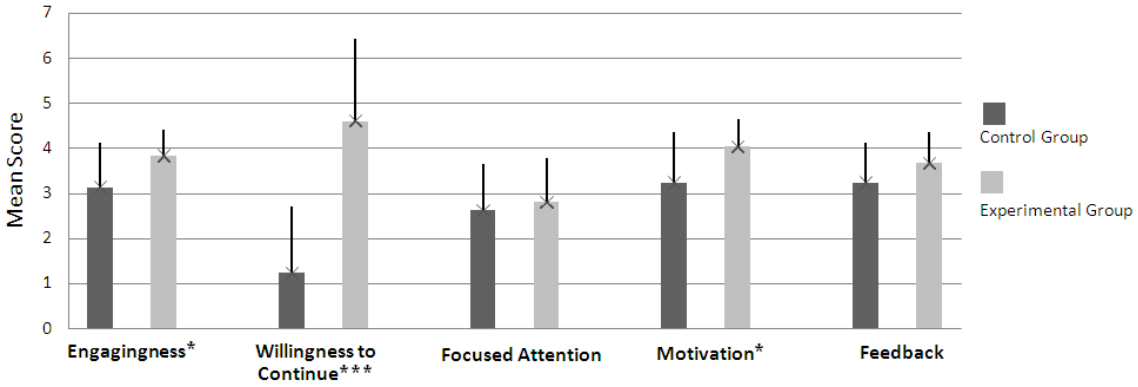


Figure 4.5: The reported scores for describing the experience of completing five training blocks of the reaching task for each participant group collected from five-point Likert scale questions. Measures showing significant differences between the two groups are indicated with the following suffix: * $p < 0.05$, *** $p < 0.001$.

The results of pair-wise comparison between the control and experimental groups of the three internally reliable descriptions of the experience measures are summarized in Table 4.7. The pair-wise comparison of the two participant groups was done using a two-tailed *t*-test and assuming unequal variances. Differences between the two groups in all three of the measures are statistically significant: the experiment group reported higher scores.

Table 4.7: Two-tailed *t*-test results for pair-wise comparison of the reported scores for the internally reliable experience measures between the two participant groups. Measures showing significant differences between the two groups are indicated with the following suffixes: * $p < 0.05$, *** $p < 0.001$.

Measure	*t*-value (*df*)	*p*-value
Engagingness*	2.13 (20)	$p = 0.04$
Willingness to Continue ***	4.73 (17)	$p < 0.001$
Motivation*	2.16 (19)	$p = 0.04$

4.3.6 Changes in Participants' Task Satisfaction and Attentiveness between the First and Last Training Blocks

The SAM questionnaire was used to gauge the changes in participants' task satisfaction and attentiveness in the duration of the exercise with the robot. Participants' in both groups reported their satisfaction and attentiveness scores after the first and the last (i.e., fifth) training blocks. The differences between the scores at the beginning and the end of the experiment were used for analysis. The mean and standard deviation of the changes in task satisfaction and attentiveness are reported in Table 4.8 and presented graphically in Figure 4.6.

Table 4.8: The means and standard deviations, in parentheses, for the changes in task satisfaction and attentiveness after five training blocks of reaching tasks. Measures showing significant differences between the two groups are indicated with the following suffix: * $p < 0.05$.

Measure	Control Group	Experimental Group
Task Satisfaction *	1.08 (1.44)	2.50 (1.72)
Task Attentiveness *	0.78 (1.59)	2.2 (1.55)

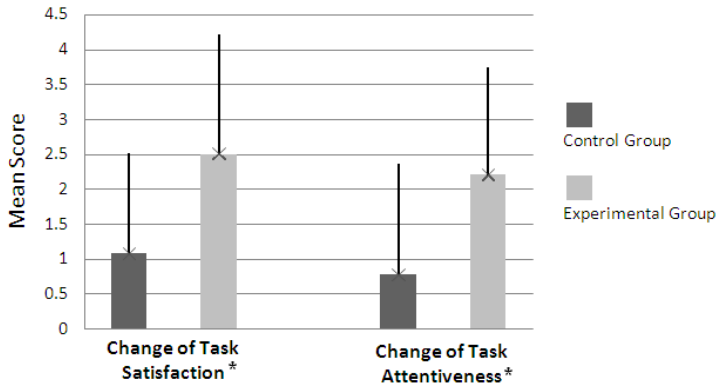


Figure 4.6: The changes in task satisfaction and attentiveness after five training blocks of reaching tasks for each participant group. Measures showing significant differences between the two groups are indicated with the following suffix: * $p < 0.05$.

A pair-wise comparison of the changes in task satisfaction and attentiveness after completing the training blocks between the two participant groups was done using two-tailed *t*-test assuming unequal variances. The control group reported statistically significant lower changes for both of the measures. Table 4.9 summarizes the results of performed *t*-tests.

Table 4.9: Two-tailed *t*-test results for pair-wise comparison of the changes in task satisfaction and attentiveness. Measures showing significant differences between the two groups are indicated with the following suffix: * $p < 0.05$.

Measure	*t*-value (*df*)	*p*-value
Task Satisfaction *	2.11(18)	0.04
Task Attentiveness *	2.17 (20)	0.04

4.4 Discussion and Conclusion

This chapter presents the results of a study in which an experimental group (participants received training blocks of a robotic reaching task at their desirable difficulty levels) was compared with a control group (order of training blocks was predefined randomly). The findings from this study demonstrate that a Neural Network model that uses a fuzzy combination of a participant's motor performance and physiological metrics can predict the direction of change to achieve the participant's desirable difficulty with an accuracy of 78% in real-time. Additionally, the results of this study show that the experimental group perceived the trio of the reaching task, the robotic manipulandum, and the experience of training with the robot as being significantly different, compared to the control group. In all the significant cases, the reported scores suggest that the experimental group's perception of the three items mentioned above is more positive than the control group.

The Neural Network (NN) model used to predict participants' desirable difficulties was trained using a data set of 107 instances. The training accuracy of this model was 89% and the closed-loop validation of it in this study yielded an accuracy rate of 78%. The training data set was built by offering the reaching task to participants at a random difficulty level in order to collect their motor performance and physiological metrics, as well as their preferred change in challenge level for the subsequent training block. Accordingly, in this study the experimental group received reaching tasks based on the output of the NN model. Although lower accuracy rate of the NN model in the closed-loop validation was expected, the high training accuracy rate of 89% can be an indicator of possible over-fitting in the training stage, an amplifying factor to deteriorate the accuracy of the predictions.

A general trend in the questionnaire data showed that practicing the reaching task under a desirable difficulty condition was less frustrating for the study participants. By practicing the task at their desirable difficulties, participants in the experimental group perceived their performance at a significantly higher level and reported lower required effort to complete the task. However, since the nature of the task (i.e., reach motions to visually presented targets within 0.5 seconds) does not change by practicing it at one's desirable difficulty, participants in both groups associated the task with the same level of mental and physical demand. Additionally, experimental group participants perceived the robot – the exercise tool – as more useful compared to the control group.

One of the main goals of the study presented in this chapter was to investigate the impacts of completing reaching tasks at one's desirable difficulty on the person's experience of the whole training session. The questionnaire data showed that participants in the experimental group were willing, on average, to continue the training session for 4.5 more training blocks. This score was significantly greater than that of the control group. Moreover, the experimental group found their experience both more engaging and more motivating compared to the control group. This is also reflected by the higher changes in the experimental group's task satisfaction and attentiveness after completing the training blocks.

The findings of this study demonstrate the ability of predicting the direction of change to reach a user's desirable difficulty in a motor learning task in real-time. This study also demonstrated the possible usefulness of practicing a task at one's desirable difficulty in improving the training experience for individuals, as well as their perception of the task and the robotic exercise tool. However, since the study used healthy participants, the effects of practice under a desirable difficulty condition on motor performance and motor learning were not studied.

Nevertheless, this study does not address how the findings of this chapter can be integrated into the field of stroke therapy and possibly address the issue of low engagement of patients during the functional recovery period. Next chapter concludes this thesis by summarizing the findings of the three human-subject studies presented, and a discussion of future work and implications of implementing the desirable difficulty methodology for the clinical population.

5 Conclusion

Therapy regimens are based on our understanding of neuroscience and motor learning principles. To stimulate patients' brain neuroplasticity and reorganization of motor pathways, therapy exercises comprise of numerous repetitions of simple physical tasks. This necessary repetitive nature of therapy exercises makes the hard work of therapy boring, which can negatively affect the patient's motivation to continue therapy.

In an attempt to address this issue, this thesis grounded its overall goals in the scientific literature of the mechanics of human engagement in tasks. According to the flow theory and the challenge point framework, completing tasks at one's desirable difficulty accelerates both learning of new skills and engagement during each exercise session. Delivering the therapy exercises at a preferred level of challenge has the potential to address the issue of "engagement in therapy regimen": Seeing one's improvement in an engaging task will increase one's self-efficacy with the task, and in the case of stroke therapy, will lead to a lower chance of therapy abandonment and faster recovery.

Any implementation of the idea of desirable difficulty engagement would rely on two main factors: having an exercise with several different difficulty levels, and offering the exercise to individuals at their desirable difficulty. These desirable difficulties depend on the user's performance and affective state during the task. While robots provide the ability to quantify the user's performance, assessing the user's affective state is more challenging and can only be done indirectly. Real-time analysis of physiological signals is one way to characterize different aspects of affect.

This thesis investigated the feasibility of implementing the idea of desirable difficulties in a robot-assisted motor rehabilitation exercise scenario to improve users' engagement in the exercise. This design process was split into three main steps: design of an effective (i.e., leading to motor learning) robotic exercise with meaningful and distinguishable levels of difficulty, development of a method to predict the user's desirable difficulty during the robotic exercise, and evaluation of the effects of exercising at one's desirable difficulty on one's perception of the exercise. The results of each step were presented in detail in Chapter 2 to Chapter 4. Overall, the main contributions of this thesis are:

1. Study I: Proposing combination of visual and force feedback error amplification as a way to promote motor adaptation, as well as demonstrating that error amplification methods can be used as a way to introduce levels of challenge into a robotic reaching task.

2. Study II: An empirical comparison of different machine learning algorithms in estimating users' preference for direction of challenge change, and proposing a fuzzy combination of motor performance and physiological metrics as predicting attributes to increase accuracy of preference estimation.

3. Study III: Providing statistical evidence that participants report a higher level of engagement with the exercise when the robot adaptively uses their affect and performance to change task difficulty.

Section 5.1 to Section 5.3 is a review of the findings of each study with a comparison with the available literature and Section 5.4 concludes this thesis by presenting recommendations for future work.

5.1 Can Error Amplification in a Robot-assisted Reaching Task Be Viewed as a Meaningful Way of Altering Task Difficulty?

The study in Chapter 2 investigated the effects of practicing reaching motions with error amplification. The focus of this study was to propose error amplification methods as a way of adding *meaningful* difficulty *levels* to a robot-assisted reaching task. *Meaningful* in this case means "leading to motor learning" and a *level* of difficulty implies that the users are able to differentiate between the amounts of challenge involved in doing a task at each of the levels.

Performance metrics were introduced to characterize healthy participants' motor adaptation to a visually distorted reaching environment (amount of improvement, speed of learning, path deviation after training). Effects of five conditions of error amplification (EA) on motor adaptation were studied. Similar to [19], the study presented in Chapter 2 could not show significant differences between adaptation properties promoted by each EA. The only exception was the high-gain visual and force feedback EA, which significantly improved the amount of improvement "*a*" in comparison to control and visual EA methods. More importantly, the main contribution of this study was to investigate the effects of EA levels on participants' affective state, an analysis that has not been performed before. Analyzing participants' questionnaire responses showed that different levels of error amplification are perceived by the study

participants as different levels of task difficulty and thus, lead to different levels of task satisfaction and attentiveness. This suggests that different levels of EA can be distinguished by users as additional challenge to the reaching task.

Several methodological differences between the Wei [18] and Celik [19] studies contributed to the fact that the ANOVA failed to show significant differences in the speed of motor adaptation measures "*b*" between EA methods. Wei and Celik used higher amplification gains (2 and 3.1) and compared error amplification using these gains with error offsetting (i.e., augmenting the cursor position by adding a constant value) in a between-subjects design. In these studies, adaptation was characterized using catch-trials (both distortion and EA are turned off: plain motion) placed between exercise-with-challenge trials that ran for over 50 cycles. This study protocol design is well suited for studying only motor adaptation. In the study presented in this chapter, the low resolution of the robot's display (640 × 480) prevented using higher gains, such as 2 and 3.1, that were validated in other studies. The study reported here intentionally avoided using a between-subjects design with long and extended exercise blocks to simulate a condition close to an actual therapy session; a participant experiences different tasks in rather short periods of time.

5.2 Can a User's Desirable Difficulty Be Predicted?

The study presented in Chapter 3 investigated methods of predicting the direction of change to reach a participant's desirable difficulty and the most informative inputs for this process in an open-loop validation process. The use of K-Nearest Neighbour, Neural Networks, and Discriminant Analysis methods was proposed as possible means of predicting users' preference. This study also explored the effectiveness of using participants' motor performance and physiological signals during the reaching task in prediction of users' desirable difficulties.

The results of this open-loop validation (k-fold cross-validation) showed that using a Neural Networks algorithm and a fuzzy combination of the motor performance and physiological metrics leads to a high preference prediction accuracy of 85%. In a follow-up study presented in Chapter 4, the same method of prediction was validated in a closed-loop setting and yielded an accuracy rate of 78%. These results are comparable to and more definitive relative to (standard deviation of prediction accuracies can be calculated in k-fold cross-validation) the available literature.

Efficiency of several machine learning methods in predicting affect based on physiological signals is presented in [55]. In that study, Rani et al. used a human-robot interaction task to elicit different affective states. In a systematic comparison of the weaknesses and strengths of four machine learning algorithms, K-Nearest Neighbour, Bayesian Network, Regression Tree, and Support Vector Machine, the study showed that all the methods performed competitively (~80% prediction accuracy). However their study used leave-one-out cross-validation, a less objective estimate of accuracy and generalizability compared to the k-fold cross-validation method used in the study of Chapter 3.

Novak et al. [58] used performance metrics and physiological measurements of users to design a biocooperative stroke rehabilitation system. In the study, Novak used different flavours of Discriminant Analysis to design a system that can predict a therapist's recommended changes for task difficulties (a more objective measure in comparison with directly asking the study participants). Their work demonstrated greater performance of adaptive algorithms that can adjust their decision making process for each individual user compared to algorithms that work based on a generic model of the users (open-loop validation 85% vs. 82%). As their study, similar to [55], used leave-one-out cross-validation, it is not possible to comment on the significance of the differences between the two prediction accuracy values. By investigation of the usefulness of each of the physiological features after the open-loop validation and exclusion of the non-relevant features, Novak designed a system that demonstrated a closed-loop prediction accuracy of 88%.

5.3 How Is a User's Experience of the Robotic Exercise Altered when the Robot Adjusts the Exercise Based on Its Prediction of the User's Challenge Preference?

The studies presented in Chapters 2 and 3 provide strong empirical evidence to support the implementation of delivering exercise at users' desirable difficulties through a robotic reaching task. The study presented in Chapter 4 investigated whether a user reports a more positive perception of the robot and the physical exercise when the robot delivers the reaching exercise at the user's desirable difficulty. Several previous studies have looked at ways of predicting a user's preference and the success of changing the behaviour of the robot based on such predictions [37, 56, 58]. However, this is the first study, to the author's knowledge, that has investigated the effects of interacting with a robot under "preferred" conditions.

In the study presented in Chapter 4, participants in a control group received training blocks with EAs in a predefined random order, while participants in an experimental group received the training blocks based on predictions of the direction of change to achieve the participants' desirable difficulty performed by a machine learning algorithm. This study focused on three items of the users' perception: 1) users' perceptions of the reaching task and the task load, 2) users' descriptions of the robot's responsiveness and usefulness as a trainer, and finally 3) users' descriptions of their experience of the entire training session. The study participants' questionnaire responses were used to quantify these items.

The experimental group participants perceived their performance at a significantly higher level and reported lower required effort to complete the task. Additionally, they perceived the robot – the exercise tool – as more useful than the control group participants did. A general trend in the questionnaire data showed that practicing the reaching task under desirable difficulty condition was less frustrating for the study participants. Moreover, the questionnaire data showed that participants in the experimental group were willing to continue the training session for approximately four times longer than the control group. Likewise, the experimental group found their experience both more engaging and more motivating compared to the control group. The higher changes in the experimental group's task satisfaction and attentiveness after completing the training blocks also imply the positive effects of exercising at one's desirable difficulties. Finally, a general trend in the questionnaire data showed that practicing the reaching task under desirable difficulty conditions was less frustrating for the study participants.

5.4 Recommendations and Future Work

The findings presented in this thesis are results of human-subject studies with healthy participants (i.e., not stroke survivors performing therapy exercises), an acceptable case for methodology validation studies. Because of this, one needs to be cautious and thoughtful in interpreting the results and generalizing them to clinical populations.

Chapter 2 of this thesis suggests a framework for adding meaningful levels of difficulty to a robotic exercise of reaching motions. The findings of this thesis about the effects of practicing with different error amplification conditions in the healthy population agree with the literature. Studies have shown that use of error amplification improves motor learning in both healthy persons and stroke survivors. As the general definition and perception of physical challenge is not affected by stroke, it is plausible to hypothesize that stroke survivors will also recognize

different levels of error amplification as different levels of reaching task challenge. However, the error amplification gains used in this study might not be suitable for the clinical population and can be too challenging cognitively and physically. Further studies are required to determine visual and/or force feedback error amplification gains that lead to better motor performance characteristics as well as a more distinctive differentiation of challenge perception in stroke population.

Chapter 3 demonstrated a possible way of predicting the direction of change to achieve a user's desirable difficulty in real-time. The method presented relies on the use of motor performance and physiological metrics as input data. The motor performance metrics introduced in this thesis characterize motor adaptation to visually distorted environments. Motor adaptation in the healthy population can be considered as a very specific case of motor learning; however, it does not alter brain motor control pathways. Because of this, motor *adaptation* happens in a shorter period of time compared to motor *learning* in the stroke population. To implement this method of user preference prediction for the stroke population, motor performance metrics suitable to characterize the slow process of motor re-learning (e.g., changes in a patient's Fugl-Meyer score over the course of treatment) must be used. Additionally, due to damage caused by stroke to the central nervous system, physiological responses are different than responses in the healthy population. Additional comparative studies between healthy and clinical populations such a Novak's work [35] may still be required to introduce physiological metrics that best capture the changes of affective states in stroke population. Based on the results of a study by Novak [58], it is highly recommended to investigate the usefulness of each of the physiological features and exclude non-relevant features in order to increase the accuracy of desired difficulty predictions.

Finally, Chapter 4 of this thesis demonstrated the possible usefulness of practicing a task at one's desirable difficulty in improving the training experience for participants, as well as improving their perception of the task and the robotic exercise tool. Physical therapy comprises of numerous training sessions; a continuous process of engagement in physical exercise, dis-engagement at the end of each exercise session, and re-engagement at the beginning of the next exercise session. The study presented in Chapter 4 was mainly focused on the user's engagement during one training session and the questionnaire items used in this study related to how dis-engagement can be delayed. To move this work to the clinical environment, it is important to study the impacts of

practice under desirable difficulties on the process of re-engagement in therapy tasks and patients' self-efficacy.

In the study presented in Chapter 4, the presented direction of change in difficulty for the control group matched the reported value by the participants (desired direction of change) in 53% of the instances. This number was 78% for the experimental group. This difference led to a significant difference in the participants' experience between the two groups. However, the study presented in Chapter 4 does not investigate the minimum threshold of participant preference prediction that leads to participant satisfaction. This will be an important factor in future work with stroke population. In this study participants finished five training blocks and it was shown that one wrong prediction in five predictions is not annoying or frustrating to them. A stroke survivor will need to complete thousands of repetitions, and with an ~80% prediction accuracy will receive over 200 training blocks in the wrong direction of difficulty change, a possibly frustrating condition. Future studies need to consider the accumulative effects of such consecutive non-desired difficulty changes.

The main goal of physical therapy for stroke survivors is to improve motor ability. Nevertheless, the repetitive nature of therapy exercises compromises patients' motivation to continue their therapy regimens. This thesis demonstrated the possibility of increasing task engagement in a typical physical therapy exercise by delivering the exercise at the user's desirable difficulty. However, since the study used healthy participants, the effects of practice at one's desirable difficulty on motor performance and motor learning were not studied. The improved engagement during an exercise session can make re-engagement in the exercise more probable, leading to an increase in the amount of time spent in active therapy. As the amount of time spent in active therapy is a predictor of both speed and amount of recovery, a hypothesis will be that delivering therapy exercises at a patient's desirable difficulty will positively impact patient's recovery of motor abilities.

References

[1] The Heart and Stroke Foundation's website, www.heartandstroke.com, accessed 21/02/2013.

[2] "AHA statistical update: heart disease and stroke statistics 2012 update –a report from the American Heart Association," Circulation, 2012, Vol. 125, pp. 2-220.

[3] J. Fang, K.M. Shaw, and M.G. George, "Prevalence of stroke - United States - 2006-2010," Morbidity & Mortality Weekly Report, Vol. 61, No. 20, 2012, pp. 379-382.

[4] J. Desrosiers, F. Malouin, D. Bourbonnais, C.L. Richards, A. Rochette, and G. Bravo, "Arm and leg impairments and disabilities after stroke rehabilitation: relation to handicap," Clinical Rehabilitation, Vol. 17, 2003, pp. 666-673.

[5] W.J. Coster, S.M. Haley, P.L. Andres, L.H. Ludlow, T.L.Y. Bond, and P.S. Ni, "Refining the conceptual basis for rehabilitation outcome measurement," Medical Care, Vol. 42, 2004, pp. 162-172.

[6] R.P. Van Peppen, G. Kwakkel, S. Wood-Dauphinee, H.J. Hendriks, P.J. VanWees, and J. Dekker, "The impact of physical therapy on functional outcomes after stroke: what's the evidence?," Clinical Rehabilitation, Vol. 18, 2004, pp. 833-862.

[7] E.M.J. Steultjens, J. Dekker, L.M. Bouter, J.C.M. van de Nes, E.H.C. Cup, and C.H.M. van den Ende, "Occupational therapy for stroke patients: a systematic review," Stroke: a Journal of Cerebral Circulation, Vol. 34, No. 3, 2003, pp. 676-687.

[8] Population Reference Bureau, "World Population Reference Sheet 2010," www.prb.org, accessed 25/10/2010.

[9] N. Hogan, and H. Krebs, "Interactive robots for neuro-rehabilitation," Restorative Neurology and Neuroscience, Vol. 22, 2004, pp. 349-358.

[10] B.R. Brewer, S.K. McDowell, and L.C. Worthen-Chaudhari, "Poststroke upper extremity rehabilitation: a review of robotic systems and clinical results," Topics in Stroke Rehabilitation, Vol. 14, 2007, pp. 22–44.

[11] M.L. Aisen, H.I. Krebs, N. Hogan, F. McDowell, and B.T. Volpe, "The effect of robot-assisted therapy and rehabilitative training on motor recovery following stroke," Archives of Neurology, Vol. 54, 1997, pp. 443-446.

[12] S. Charles, H.I. Krebs, B. Volpe, D. Lynch, and N. Hogan, "Wrist rehabilitation following stroke: initial clinical results," In Proceedings of IEEE International Conference on Rehabilitation Robotics ICORR, 2005, pp. 13-16.

[13] P.S. Lum, C.G. Burgar, P.C. Shor, M. Majmundar, and H.F.M. Van der Loos, "Robot-assisted movement training compared with conventional therapy techniques for the rehabilitation of upper-limb motor function after stroke," Archives of Physical Medicine and Rehabilitation, Vol. 83, 2002, pp. 952-959.

[14] A. Gupta, M.K. O'Malley, V. Patoglu, and C. Burgar, "Design, control and performance of RiceWrist: a force feedback wrist exoskeleton for rehabilitation and training," The International Journal of Robotics Research, Vol. 27, 2008, pp. 233-251.

[15] F. Amirabdollahian, R. Loureiro, E. Gradwell, C. Collin, W. Harwin, and G. Johnson, "Multivariate analysis of the Fugl-Meyer outcome measures assessing the effectiveness of GENTLE/S robot-mediated stroke therapy," Journal of NeuroEngineering and Rehabilitation, Vol. 4, 2007, p. 4-19.

[16] M.J. Johnson, H.F.M. Van der Loos, C.G. Burgar, P. Shor, and L.J. Leifer, "Design and evaluation of Driver's SEAT: a car steering simulation environment for upper limb stroke therapy," Robotica, Vol. 21, 2003, pp. 13-23.

[17] J.L. Patton, M.E. Stoykov, M. Kovic, and F.A. Mussa-Ivaldi, "Evaluation of robotic training forces that either enhance or reduce error in chronic hemiparetic stroke survivors," Experimental Brain Research, Vol. 168, 2006, pp. 368-383.

[18] Y. Wei, P. Bajaj, R.A. Scheidt, and J.L. Patton, "Visual error augmentation for enhancing motor learning and rehabilitative relearning," In Proceedings of IEEE International Conference on Rehabilitation Robotics ICORR, 2005, pp. 505-510.

[19] O. Celik, D. Powell, and M.K. O'Malley, "Impact of visual error augmentation methods on task performance and motor adaptation," In Proceedings of IEEE International Conference on Rehabilitation Robotics ICORR, 2009, pp. 793-798.

[20] J.W. Burke, M.D.J. McNeill, D.K. Charles, P.J. Morrow, J.H. Crosbie, and S.M. McDonough, "Optimising engagement strategies for stroke rehabilitation using serious games," Visual Computing, Vol. 25, 2009, pp.1085-1099.

[21] P.W. Duncan, "Adherence to postacute rehabilitation guidelines is associated with functional recovery in stroke," Stroke, Vol. 33, No. 1, 2002, pp.167-178.

[22] N. Maclean, P. Pound, C. Wolfe, and a Rudd, “Qualitative analysis of stroke patients’ motivation for rehabilitation,” BMJ, Vol. 321, 2000, pp. 1051-1054.

[23] M. Kotila, H. Numminen, O. Waltimo, and M. Kaste, “Depression after stroke: results of the FINNSTROKE study,” Stroke, Vol. 29, 1998, pp.368-372.

[24] K. Oliver, and T.Cronan, “Predictors of exercise behaviors among fibromyalgia patients,” Preventive Medicine, Vol. 33, No. 4, 2002, pp. 383-389.

[25] P.S. Pohl, J.M. McDowd, D. Filion, L.G. Richards, and W. Stiers, “Implicit learning of a motor skill after mild and moderate stroke,” ClinicalRehabilitation, Vol. 20, 2006, pp. 246-253.

[26] M. Csikszentmihalyi, “Flow: The Psychology of Optimal Experience,” Harper Perennial, 1991.

[27] J. A. Russell, “A Circumplex Model of Affect,” Journal of Personality and Social Psychology, Vol. 39, No. 6, 1980, pp. 1161–1178.

[28] J. Nakamura, and M. Csikszentmihalyi, "The concept of flow," Handbook of Positive Psychology, Oxford University Press, 2005, pp. 89–105.

[29] M.A. Guadagnoli, and T.D. Lee, “Challenge point: a framework for conceptualizing the effects of various practice conditions in motor learning,” Journal of Motor Behaviour, Vol. 36, No. 2, 2004, pp. 212–224.

[30] J.A. Russell, A. Weiss, and G.A. Mendelsohn, “Affect Grid: A single-item scale of pleasure and arousal,” Journal of Personality and Social Psychology, Vol. 57, 1989, pp. 493-502.

[31] D. Kulić, and E. A. Croft, “Estimating robot induced affective state using Hidden Markov Models,” In Proceedings of IEEE International Symposium on Robot and Human Interactive Communication RO-MAN, 2006, pp. 257-262.

[32] D. Kulić, and E.A. Croft, “Experimental validation of affective state estimation and safe motion planning during human robot interaction,” Robotica, Vol. 25, 2007, pp. 13-27.

[33] M.K.X.J. Pan, J. Chang, G. H. Himmetoglu, A. Moon, T.W. Hazelton, K.E. MacLean, and E.A. Croft, “Now, where was I? physiologically-triggered bookmarking,” In Proceedings of the Annual Conference on Human Factors in Computing Systems CHI, 2011, pp. 363-372.

[34] S. Zoghbi, D. Kulić, E. Croft, and H.F.M. Van der Loos, "On line - affective state reporting device: a tool for evaluating affective state inference systems," Late Breaking Abstracts, ACMIEEE International Conference on Human-Robot Interaction HRI, 2009.

[35] D. Novak, J. Ziherl, A. Olensek, M. Milavec, J. Podobnik, M. Mihelj, and M. Munih, "Psychophysiological responses to robotic rehabilitation tasks in stroke," IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 18, No. 4, 2010, pp. 351-361.

[36] M. Munih, D. Novak, T. Bajd, and M. Mihelj, "Biocooperation in rehabilitation robotics of upper extremities," In proceedings of IEEE International Conference on Rehabilitation Robotics ICORR, 2009, pp. 425-430.

[37] C. Liu, P. Rani, and N. Sarkar, "Affective state recognition and adaptation in human-robot interaction: a design approach," In Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems IROS, 2006, pp. 3099-3106.

[38] C.M. Dean, and R.B. Shepherd, "Task-related training improves performance of seated reaching tasks after stroke. a randomized controlled trial," Stroke, Vol. 28, 1997, pp. 722-728.

[39] T.R. Kaminski, C. Bock, and A.M. Gentile, "The coordination between trunk and arm motion during pointing movements," Experimental Brain Research, Vol. 106, 1995, pp. 702-707.

[40] S.H. Scott, "Role of motor cortex in coordinating multijoint movements: is it time for a new paradigm?" Canadian Journal of Physiology and Pharmacology, Vol. 78, 2000, pp. 923-933.

[41] T. Flash, and N. Hogan, "The coordination of arm movements: an experimentally confirmed mathematical model," The Journal of Neuroscience, Vol. 5, No. 7, 1985, pp. 1688-1703.

[42] P. Morasso, "Spatial control of arm movements," Experimental Brain Research, Vol. 42, 1981, pp. 223-227.

[43] A.J. Bastian, T.A. Martin, J.G. Keating, and W.T. Thach, "Cerebellar ataxia: abnormal control of interaction torques across multiple joints," Journal of Neurophysiology, Vol. 76, 1996, pp. 492-509.

[44] D.M. Wolpert, Z. Ghahramani, and M.I. Jordan, "An internal model for sensorimotor integration," Science, Vol. 269, 1995, pp. 1880–1882.

[45] D. B. Willingham, "The neural basis of motor-skill learning," Current Directions in Psychological Science, Vol. 8, No. 6, 1999, pp. 178-182.

[46] Y. Matsuoka, B. Brewer, and R. Klatzky, "Using visual feedback distortion to alter coordinated pinching patterns for robotic rehabilitation," Journal of NeuroEngineering and Rehabilitation, Vol. 4, No. 1, 2007, pp. 17-25.

[47] B. Brewer, R. Klatzky, and Y. Matsuoka, "Initial therapeutic results of visual feedback manipulation in robotic rehabilitation," In Proceedings of International Workshop on Virtual Rehabilitation, 2006, pp. 160-166.

[48] F. Abdollahi, S. Rozario, R. Kenyon, J. Patton, E. Case, M. Kovic, and M. Listenberger, "Arm control recovery enhanced by error augmentation," In Proceedings of IEEE International Conference on Rehabilitation Robotics ICORR, 2011, pp. 1-6.

[49] W.J. Atsma, "Inference of central nervous system input and its complexity for interactive arm movement," Ph.D. dissertation, Department of Mechanical Engineering, University of British Columbia, Vancouver, BC, Canada, 2006.

[50] M. Bradley, and P. Lang, "Measuring emotion: the self-assessment manikin and the semantic differential," Journal of Behavior Therapy and Experimental Psychiatry, Vol. 25, 1994, pp. 49-59.

[51] S.A. Glantz, "Primer of bio-statistics," McGraw-Hill, 4th Edition, 1997.

[52] A. Agresti, B. Finlay, "Statistical Methods for the Social Sciences," Pearson Prentice Hall, 4th Edition, 2009.

[53] J.D. Schaechter, "Motor rehabilitation and brain plasticity after hemiparetic stroke," Progress in Neurobiology, Vol. 73, pp. 61-72, 2004.

[54] N. Shirzad, and H.F.M. Van der Loos, "Error amplification to promote motor learning and motivation in therapy robotics," In Proceedings of IEEE International Conference on Engineering in Medicine and Biology Society EMBC, 2012, pp. 3907-3910.

[55] P. Rani, C. Liu, N. Sarkar, E. Vanman, "An empirical study of machine learning techniques for affect recognition in human-robot interaction," Pattern Analysis and Applications, Vol. 9, 2006, pp. 58-69.

[56] C. Liu, K. Conn, N. Sarkar, and W. Stone, "Online affect detection and robot behavior adaptation for intervention of children with autism," IEEE Transactions on Robotics, Vol. 24, No. 4, 2008, pp. 883-896.

[57] F. Wang, D.E. Barkana, and N. Sarkar, "Impact of visual error augmentation when integrated with assist-as-needed training method in robot-assisted rehabilitation," IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 18, No. 5, 2010, pp. 571-579.

[58] D. Novak, M. Mihelj, J. Ziherl, A. Olensek, and M. Munih, "Psychophysiological measurements in a biocooperative feedback loop for upper extremity rehabilitation,"IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 19, No. 4, 2011, pp. 400-410.

[59] S. Russell, P. Norvig, "Artificial intelligence: a modern approach," Pearson Prentice Hall, 3rd Edition, 2010.

[60] M.T. Hagan, H.B. Demuth, M.H. Beale, "Neural network Design," Campus Publishing Service, Colorado University Bookstore, 2002.

[61] S.G. Hart, and L.E. Staveland, "Development of NASA-TLX (task load index): results of empirical and theoretical research," In P. A. Hancock and N. Meshkati (Eds.) Human Mental Workload, North Holland Press, 1988.

[62] S.G. Hart, "NASA-task load index (NASA-TLX); 20 years later," In Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting, 2006, pp. 904-908.

[63] C. Bartneck, D. Kulić, E.A. Croft, and S. Zoghbi, "Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots," International Journal of Social Robotics, Vol. 1, No. 1, 2008, pp. 71-81.

[64] H.L. O'Brien, and E.G. Toms, "What is user engagement?," Journal of the American Society for Information Science Technology, Vol. 59, No. 6, 2008, pp. 938-955.

[65] H.L. O'Brien, and E.G. Toms, "The development and evaluation of a survey to measure user engagement," Journal of the American Society for Information Science , Vol. 61, No. 1, 2010, pp. 50-69.

Appendix A – Control Architecture of the Robotic Manipulandum

This appendix presents the details of the control architecture of the robotic manipulandum used in all of the three studies of this thesis. Section A.1 briefly reviews the specifications of the robotic manipulandum. Section A.2 presents how reaching trajectory errors were calculated and how error amplification methods were implemented in the control architecture of the robot.

A.1 Overview of the Robotic Manipulandum

The robotic manipulandum that was used in this thesis is a five bar robot that was initially developed by Atsma [49] in the Neuromotor Control Laboratory at the University of British Columbia to investigate different aspects of motor learning. This robot has two degrees of freedom in the horizontal plane. Two direct-drive motors (Parker-Compumotor Dynaserv DR1060B) located at the base run the robot and the two "elbow" joints of the robot are passive (Figure A.1). Encoders integrated with the motors supply position feedback while a multi-axes force sensor (ATI Industrial Automation Inc. "Mini" sensor) at the robot's end-effector is used to measure the interaction forces in the horizontal plane. The robot's end-effector is also instrumented with a handle: Users hold on to this handle and move the robot around.

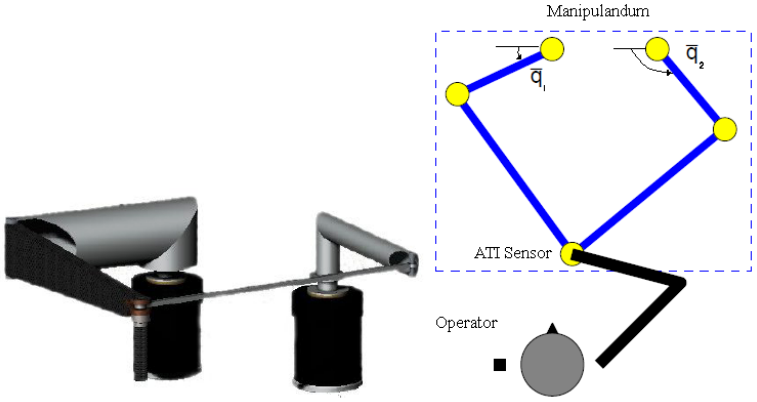


Figure A.1: The 5 bar manipulandum on left. The diagram on the right shows how an operator interacts with it. This figure was originally used in [49] and has been used with permission.

Figure A.2 shows the hardware configuration of this robotic manipulandum. The controller program is run on the main computer workstation which hosts a Digital Signal Processor (DSP) board (dSpace Inc. DS1102). This DSP board also provides data logging facilities for the MATLAB environment. The DSP acts as a communication port between the main computer and the robot's motors. It receives position feedback from the motor drives and communicates torque commands. A second computer runs A program called TargetDisplay (TD) which provides

visual feedback to the operator on a flat screen monitor. TD is also used to read the force data from the ATI sensor. TargetDisplay communicates with the DSP board on the first computer through a serial port.

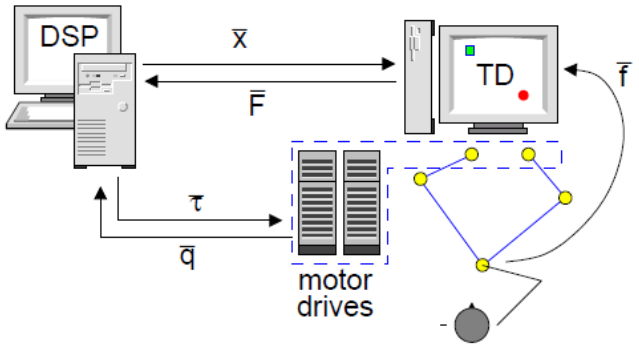


Figure A.2: Schematic of the control system hardware. This figure was originally used in [49] and has been used with permission.

Figure A.3 presents the original control scheme of the manipulandum. This impedance controller outputs a toque command to each of the two motors (Figure A.3 top right) based on the force that the operator is applying to the end effector (user's desired motion measured by the TargetDisplay interface), the predesigned perturbation criteria that were used in studies conducted by Atsma [49], and the position of the end effector ("bounding box" box in the model the ensures the operator stays within the working area). Atsma used velocity-based perturbation criteria in his studies. This was modified to implement the position-based perturbation criterion (force feedback error amplification) of this thesis. The friction compensation box on the bottom right portion of Figure A.3 compensates for high friction and cogging torque in one of the motors.

A.2 Calculation of Trajectory Error and Implementation of Error Amplification

In the studies presented in this thesis, participants were instructed to move a cursor to target points presented to them on a monitor using the robot. The ideal reaching path would be a line between the starting point of the motion and the target point. Lateral deviation from this line at each time during the reaching motion was called reaching trajectory error and would be augmented via visual means or force feedback to implement the idea of error amplification.

To amplify reaching trajectory errors, the error vector must be defined. This was done using simple vector calculus demonstrated by Figure A.4. In Figure A.4, start point *a*, target point *b*, and the current cursor position *c* are defined. Since these points are known the ideal path vector

ab and the cursor motion vector *ac* can be defined. Then the goal is to find the magnitude and direction of the error vector *cd*.

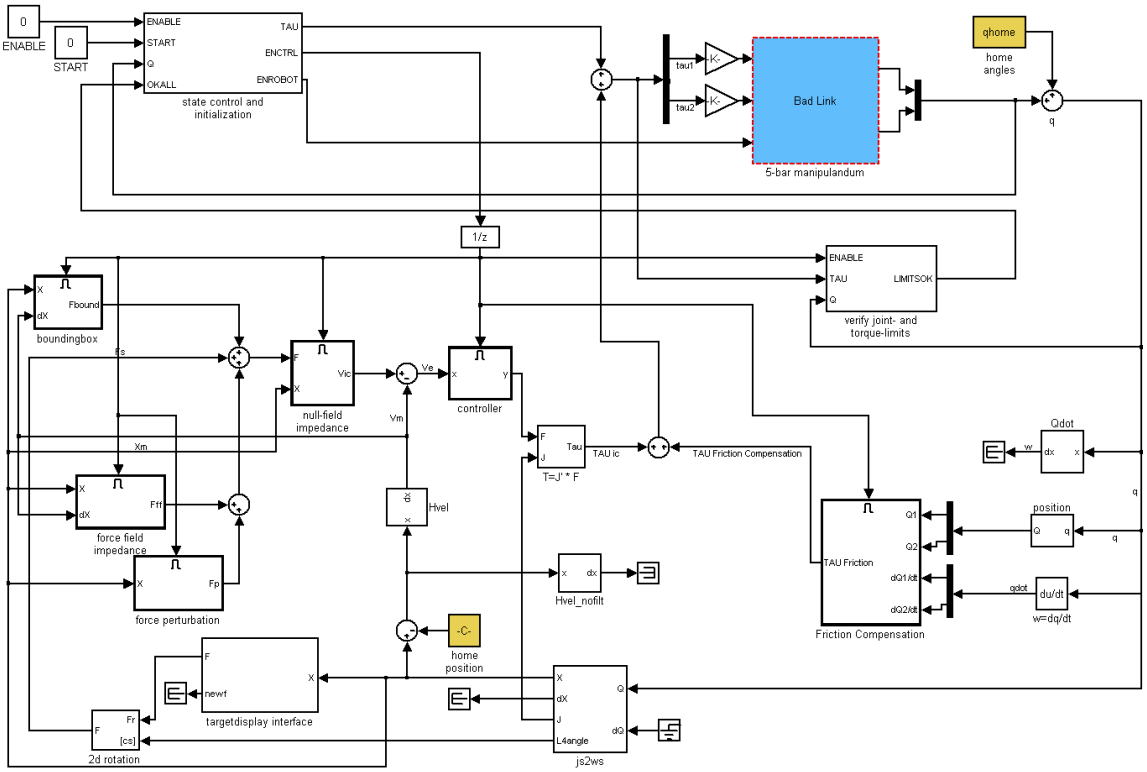


Figure A.3: Screen Capture of the original impedance controller of the manipulandum implemented in MATLAB Simulink by Atsma [49].

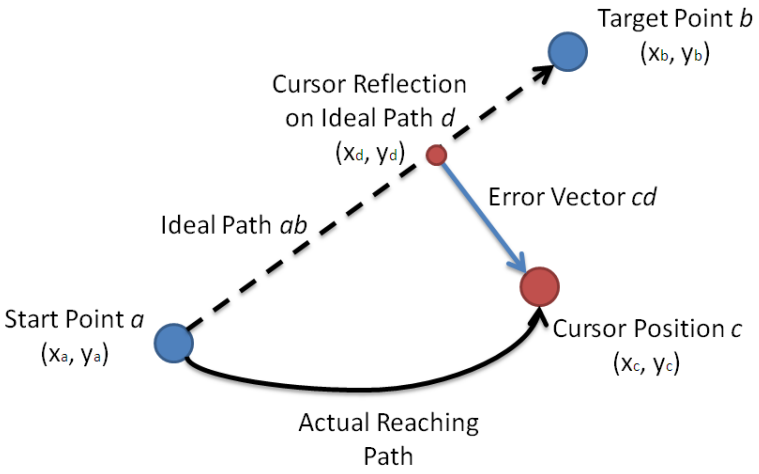


Figure A.4: Vector calculus to calculate the reaching trajectory error vector

Trajectory error is defined as lateral deviation from the ideal path. This means that the error vector is perpendicular to the ideal path. Then the normal vector of the error (direction of error) can be defined by the following equation:

$$n_{cd} = \frac{(a_y - b_y, b_x - a_x)}{|ab|} * sign(ab \times ac) \quad \text{(A.1)}$$

Forming a triangle between points *d*, *a*, and *c*:

$$\sin(\widehat{dac}) = \frac{|cd|}{|ac|} \quad \text{(A.2)}$$

Also:

$$\sin(\widehat{dac}) = \frac{|ab \times ac|}{|ab| * |ac|} \quad \text{(A.3)}$$

Combining Equation A.2 and A.3 gives the error vector's magnitude:

$$|cd| = \frac{|ab \times ac|}{|ab|} \quad \text{(A.4)}$$

Figure A.5 shows a conceptual representation of the control scheme that was implemented for the studies of this thesis. The same way of error vector calculation presented by Equation A.1 to Equation A.4 was used in the robot's control architecture, depicted as "Error Calc." box in Figure A.5. The force sensor on the robot's handle provides the user's direction of motion and encoders at the robot's motors measure the rotation of motors (q or joint coordinates). Using forward kinematics the Jacobian matrix and end-effector position x are calculated. In the control condition (no error amplification), the impedance controller designed by Atsma [49] was used to present x on the screen as the cursor position and to calculate the required motor torques based on the force sensor readings.

End-effector, start and target point coordinates were used to calculate the error vector. To implement visual error augmentation, vector summation of end-effector position vector x and the amplified error vector (by a gain of α) was calculated and this augmented end-effector position was presented on the monitor as the cursor position.

Visual and force feedback error amplification calculations were very similar. The augmented end-effector position was used to visually amplify the error vector. In addition to this vector, an augmented force vector was calculated as the scalar product of an amplification gain and the error vector. This augmented force vector was added to the sensor reading and was used to calculate the required motor torques.

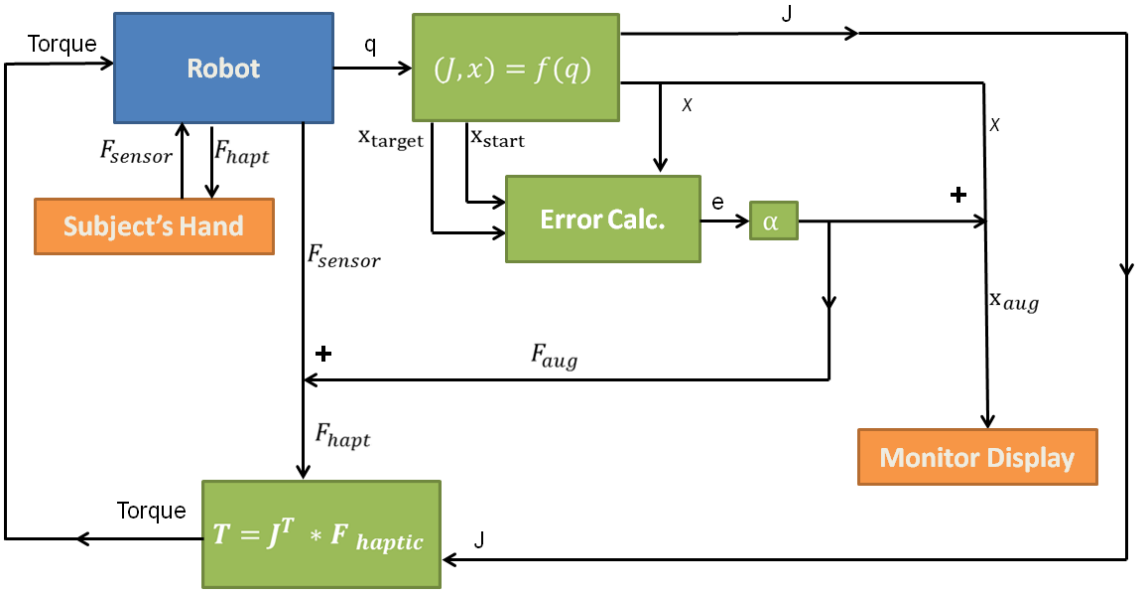


Figure A.5: Conceptual Demonstration of the Robot's Control Architecture to Implement Error Amplification Methods.

Appendix B – Advertisements, Consents, and Questionnaires

This appendix presents the details of the study advertisements used to recruit participants, consent forms, and the questionnaires used in the human-subject studies of this thesis.

B.1 Study Advertisements

Figure B.1 presents the "call for volunteers" that was posted around the campus of the University of British Columbia to recruit participants for the study presented in Chapter 2 of this thesis. Figure B.2 presents the "call for volunteers" that was posted around the campus of the University of British Columbia to recruit participants for the two studies presented in Chapter 3 and Chapter 4 of this thesis.

THE UNIVERSITY OF BRITISH COLUMBIA

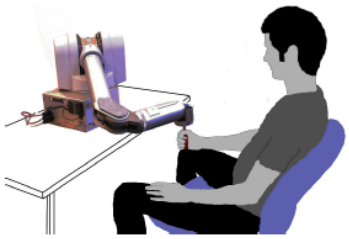
Department of Mechanical Engineering
The University of British Columbia
2054-6250 Applied Science Lane
Vancouver, B.C. Canada V6T 1Z4
Tel: (604) 822-2781
Fax: (604) 822-2403

Call for Volunteers
Robotic Stroke Therapy Study

We are seeking volunteers to help contribute to the development of a robotic therapy system to be used in stroke rehabilitation.

This research is designed to investigate the effectiveness of use of a person's physiological signals in making robotic stroke therapy more sustainable. The final outcome of this research will be a robotic system capable of modulating rehabilitation training tasks based on a person's performance and feelings.

The experiments will be carried out in the Collaborative Advanced Robotics and Intelligent Systems (CARIS) Laboratory, Room X015, ICICS building, 2366 Main Mall at the University of British Columbia, Vancouver. Participants will be asked to complete a series of reaching tasks in a virtual environment using a robotic system. The total time commitment will be approximately 60-75 minutes. Participants will be required to complete a consent form before participating in this study. **We are seeking healthy volunteers with no reported neurological deficit.**



For information/concerns regarding this study please contact:
Navid Shirzad
<omit>
<omit>

Navid Shirzad, Masters Candidate, UBC Mechanical Engineering, <omit>
Mike Van der Loos, Associate Professor, UBC Mechanical Engineering, vdl@mech.ubc.ca

Robotic Stroke Therapy Study <omit> <omit>	Robotic Stroke Therapy Study <omit> <omit>	Robotic Stroke Therapy Study <omit> <omit>	Robotic Stroke Therapy Study <omit> <omit>	Robotic Stroke Therapy Study <omit> <omit>	Robotic Stroke Therapy Study <omit> <omit>	Robotic Stroke Therapy Study <omit> <omit>	Robotic Stroke Therapy Study <omit> <omit>	Robotic Stroke Therapy Study <omit> <omit>

Version 3 (July 25, 2011)

Figure B.1: Contents of the "call for volunteers" that was posted around the campus of the University of British Columbia to recruit participants for the study presented in Chapter 2.

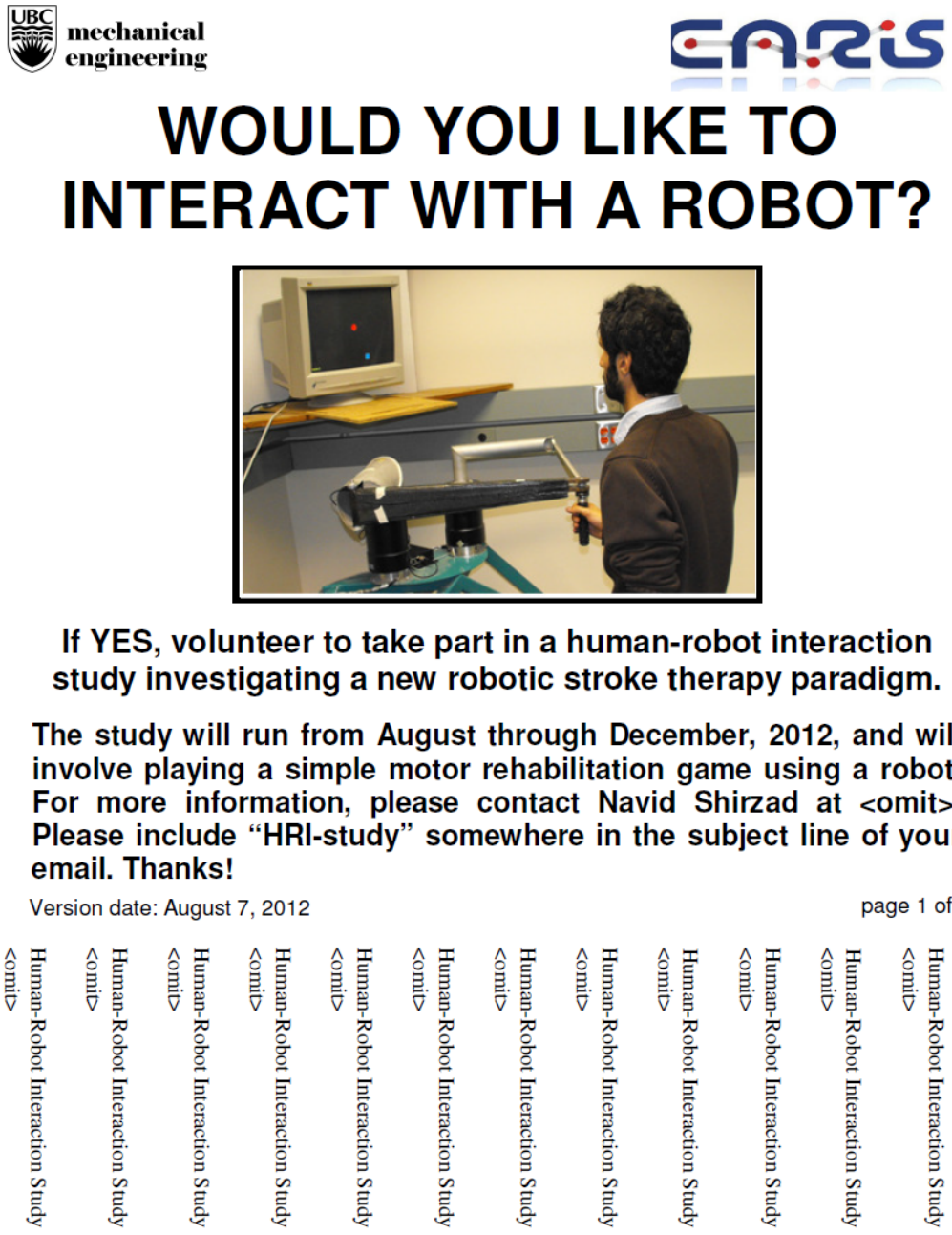


Figure B.2: Contents of the "call for volunteers" that was posted around the campus of the University of British Columbia to recruit participants for the studies presented in Chapter 3 and Chapter 4.

B.2 Information Booklet and Consent Form

The information booklet and consent form presented to the study participants were same for all of the three studies presented in this thesis. Figure B.3 to Figure B.8 presents the contents of the information and consent form booklet.

THE UNIVERSITY OF BRITISH COLUMBIA

Department of Mechanical Engineering
The University of British Columbia
2054-6250 Applied Science Lane
Vancouver, B.C. Canada V6T 1Z4
Tel: (604) 822-2781
Fax: (604) 822-2403

Informed Consent Form

Distortion-based interface design for sustainable robotic therapy

Principal Investigator:
Dr. H.F. Machiel (Mike) Van der Loos, Ph.D., P.Eng., Associate Professor, Department of Mechanical Engineering, University of British Columbia. Phone: (604) 827-4479

Co-Investigator:
Navid Shirzad, B.Sc., Department of Mechanical Engineering, University of British Columbia. Phone: <omit> or <omit>.

Version 4 (July 25, 2011)

Figure B.3: Information and consent form booklet used in all of the three studies presented in this thesis (page 1 of 6).

Subjects Participating in the Study:
We are developing a new robotic system for upper body movement therapy. We are recruiting able bodied, English speaking adults. If you have any of these conditions, you will not be able to participate in the study:

- A diagnosis for any neurological impairment
- Any condition that limits the use of your arm
- Mini-mental status exam score less than 24
- Legally blind
- Under current rehabilitation care
- An upper-limb amputation
- Uncorrected poor vision
- Cannot use the adjustable chair used for the experiment
- Cannot personally give informed consent

You may withdraw from the experiment at any time by contacting Navid Shirzad at <omit>. You do not need to provide a reason for your withdrawal if you do not wish to participate.

Purpose:
This study is exploring new ways to do physical therapy for your arm using a robot. Use of robots in stroke therapy is becoming more popular, since robots can give people many different kinds of exercises. The difficulty of the exercises can be changed in many ways. For example, your heart rate and breathing rate, which are two common physiological signals, can be measured and used to change the robot's movements. Another type of physiological signal is skin conductance, which measures how sweaty your skin is. All of these signals can be measured using small, stick-on sensors. In this way, your physiological signals can be used to make the exercises more or less difficult.

Recent studies have shown that your feelings have a direct relation to your physiological signals. So, based on sensor readings, our software can interpret your feelings to some extent. For example, the sensors can tell us if you are more or less happy, or how active you are feeling. The software will then use this information to change the difficulty of the exercises. The goal is to keep you engaged with the therapy task and not make it too hard or too easy. In this way, we can tune the exercises to your changing feelings. Our goal is to show that this can result in a better stroke therapy program.

The physical exercise itself consists of a series of arm reaching motions. The end of the robot has a handle and you will be pushing and pulling on the handle according to instructions shown on the screen.

In this phase of the study, we are recruiting 40 healthy control subjects. In later phases of this study, we will recruit 30 stroke survivor subjects.

Figure B.4: Information and consent form booklet used in all of the three studies presented in this thesis (page 2 of 6).

Study Procedures:
We will conduct the experiments in the Collaborative Advanced Robotics and Intelligent Systems (CARIS) Laboratory. The CARIS Laboratory is in room X015 of ICICS building, 2366 Main Mall, UBC Vancouver campus.

The experiment consists of one session of 60-75 minutes total. The time for setup is around 15-20 minutes. Then the experiments will run 45-55 minutes. We will photograph and video record the experiment for analysis purposes only. We will place all the recorded data on a password-protected computer in the CARIS Laboratory.

We will assign each participant a number and this participant number will be the only mean to identify all data files and written records. We are intending to present the data at scientific conferences and peer-reviewed journals. We will present the results as mean values over the group of participants. To present individual data, we will refer to it by participant number.

After completion of the study, all data will be stored on CD/DVD discs. We will keep the discs in a locked cabinet in the CARIS Laboratory. In addition, we will use the data only in relation to this specific study. We will mechanically destroy the data after a maximum of five years.

We will begin the experiment by introducing the different pieces of equipment used in this study. We will show you the robot arm used for the exercises and the sensors we use to measure your physiological signals.

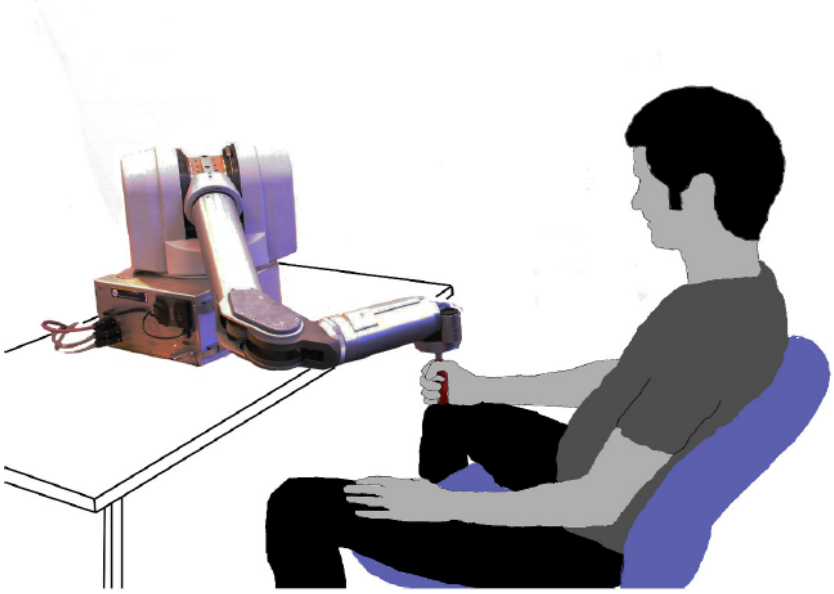


Figure : a human subject interacting with the robotic arm

Figure B.5: Information and consent form booklet used in all of the three studies presented in this thesis (page 3 of 6).

The sensors are attached to your skin by small adhesive strips. We will sterilize the sensors before we put them on your skin. We will use three types of sensors. The temperature sensor will be on a strap around your one finger. The breathing rate sensor will be on a strap around your chest. The skin conductance sensor will be on two small straps around two fingertips.

After this, we will seat you in front of the robotic arm. We will train you on how to use the robotic arm to complete the tasks. We will instruct you on safety features of the device and how to move the robot. You will perform some practice trials to make yourself familiar with the system and tasks.

The experiment will consist of a number of blocks of reaching tasks. In these reaching tasks, we ask you to move a cursor to a target point. Both the cursor and target will be visually presented on a monitor. The handle on the end of the robot allows you to move the cursor on the screen. We will ask you to reach to the target as fast and as accurately as possible. You should be able to do each of the reaching tasks in about one second.

We will sometimes add a sideways force (of no more than 2 pounds) to the robot as you are reaching for the target, or distort the way the cursor is being displayed to you. Sideways forces and visual distortion may make the reaching task harder to do.

We will measure your physiological data as you do the tasks. This consists of skin conductance response, heart rate and breathing rate. We will use the data to tune the robotic system and experimental task.

Questionnaire:
Upon completion of each block of tasks, we will ask you to fill in a questionnaire. The name of this questionnaire is the Self-Assessment Manikin (SAM). We use a SAM questionnaire to gauge your feelings about the experience of doing each block of tasks. The questions relate to how much you liked the trials, how attentive you were and how in control you felt. We will use the data from this questionnaire to assess your overall level of engagement in the task. We will not require you to answer questions that you do not feel comfortable answering.

Risks:

1) The force you need to move the robot arm is low. However, you may experience some physical fatigue. You may feel some mental fatigue as well. Paying attention to the tasks can cause mental fatigue. There will be breaks scheduled between blocks of tasks. You will be able to rest at any time if you ask the experimenter for a pause.
2) This study involves the development of novel therapy strategies. In therapy, we aim to promote behaviour change. Here, behaviour refers to the way participants perform reaching tasks. Former studies have involved a robotic reaching task within a distorted virtual environment. Distortion refers to use of deliberately inaccurate feedback. Visual exaggeration of cursor position error is an example of distortion. People easily adapt to the distortion over a number of trials. In healthy subjects, upon removing distortion, this adaptation quickly reverses. We call this process "washout". Because of washout, there is no residual effect of distortion in healthy subjects.

Figure B.6: Information and consent form booklet used in all of the three studies presented in this thesis (page 4 of 6).

On the other hand, for stroke survivors, adaptation effects appear to be more permanent when performed over a large number of trial sessions.

Our study consists of a session of small number of trials. This is much less than the thousands of trials needed to demonstrate a therapeutic effect. Thus, in our study, there is very low risk of any permanent changes in performance.

This study examines effect of distortion on your engagement with the task. In later studies, we will combine the knowledge we have gained here with intensive therapy sessions in a longer-term clinical trial. We expect more permanent performance changes in such a clinical trial.

Benefits for participation in this study:
You will not receive any financial benefits from participating in this study. You can learn about the ability of the human brain to adapt to physical inputs.

Confidentiality:
We will keep all personal information confidential. We will keep all documents in a locked filing cabinet in the principal investigator's laboratory. We will store digital data, including photographs and video records, on password-protected computers. Only the investigators involved in this study will have access to these computers. We will refer to individual data in reports or scientific publications by subject number only.

We will blur identifying features in video and photographs presented in publications. Access to the photographs and video will be restricted to the investigators. We will use the collected data only in relation to this particular study. We will keep data obtained from this study for at least 5 years after the dissemination of results. After this period, we will destroy the data to assure anonymity.

Signing this consent form will not limit your legal rights against the sponsor, investigators, or anyone else. We will respect your confidentiality. We will not release any information that discloses your identity without your specific consent. We do not allow any records that identify you by name or initials to leave the investigator's offices.

Payment:
You will not receive any payment for your participation. However, we will reimburse public transit expenses or provide free parking. We require receipts to be able to reimburse you.

Contact:
If you have any questions about this study, you may contact Dr. Mike Van der Loos at (604) 827-4479 or Navid Shirzad at <omit>.

If you have any concerns about your treatment or rights as a research subject, you may contact the Research Subject Information Line at the University of British Columbia at (604) 822-8598.

Figure B.7: Information and consent form booklet used in all of the three studies presented in this thesis (page 5 of 6).

Subject's Consent to Participate:
I understand that my participation in this study is voluntary and that I may refuse to participate or I may withdraw from the study at any time without consequences.

Signing this consent form does not limit my legal rights against the sponsor, investigators, or anyone else.

I have had sufficient time to consider the information provided and to ask for advice if necessary.

I will receive a signed and dated copy of this consent form for my own records.

I understand that I will not be paid for my participation in this study.

I consent to participate in this study by signing in the space provided below.

Subject's Signature	**Printed name of subject**	Date
Witness Signature	**Printed name of witness**	Date
Investigator's Signature	**Printed name of Principal Investigator**	Date

Version 4 (July 25, 2011)

Figure B.8: Information and consent form booklet used in all of the three studies presented in this thesis (page 6 of 6).

B.3 Questionnaires

In all the three studies presented in this thesis, participants were required to score higher than 24 on a Follestini Mini-Mental test. Figure B.9 to Figure B.11 demonstrates contents of this test. The study participants were asked to report their level of satisfaction, attentiveness, and perceived control over the robotic manipulandum in the studies presented in Chapter 2 and Chapter 4 of this thesis. The Self Assessment Manikin questionnaire was used for this purpose (Figure B.12 and Figure B.13). The post-experiment questionnaire used in the study presented in Chapter 4 is demonstrated by Figure B.14 to Figure B.17.

FOLSTEIN MINI – MENTAL STATE TEST

BEFORE YOU START: IF THE MINI-MENTAL STATUS IS NOT PERFORMED SPECIFY THE REASON BELOW

01 APHASIA (DO 3-STEP COMMAND)
02 TRACHEOSTOMY
03 OTHER (SPECIFY) _____
99 REFUSED

ORIENTATION:

Name the: Year, Month, Day, Date, Season _____ / 5
Name the: State, City, Country, Hospital (Building), Floor _____ / 5

REGISTRATION:

Examiner names 3 objects (e.g., apple, table, penny).
Ask the patient/subject to repeat the names of the three items.
Score one for each correct answer. _____ / 3

Ask the patient/subject to now remember the names of the three items you used in the above exercise (repeat them with the patient/subject until correct).
Inform the patient/subject that they need to remember these objects, and they will be asked to recall the names of these objects later.

ATTENTION AND CALCULATION:

Ask the patient/subject to subtract 7 from 100, then subtract 7 from the result, etc.
Have the patient/subject continue to subtract 7 from the result a total of 5 times.
(100, 93, 86, 79, 72, 65)

OR

Ask the patient/subject to spell the word "world" backwards. _____ /5

RECALL:

Ask the names of the three items learned earlier. _____ / 3

LANGUAGE:

Ask the patient/subject to name a "pencil" and a "watch". _____ / 2

Ask the patient/subject to repeat: "No ifs, ands or buts." _____ / 1

Version 1 (July 7, 2011)

Figure B.9: Contents of the Follestini Mini-Mental Test used in all three of the studies of this thesis (page 1 of 3).

Give the patient/subject a three stage command.
Score one for each stage (e.g., Take the piece of paper,
fold the paper in half, hand it back to me) _____ / 3

Ask the patient/subject to read and obey the written command on the piece of paper in front of them. (The command is "Close your eyes"). _____ / 1

Ask the patient/subject to write a sentence.
If the sentence is sensible and has a subject and a verb score one point. _____ / 1

COPY:

Ask the patient/subject to copy a pair of intersecting pentagons. _____ / 1

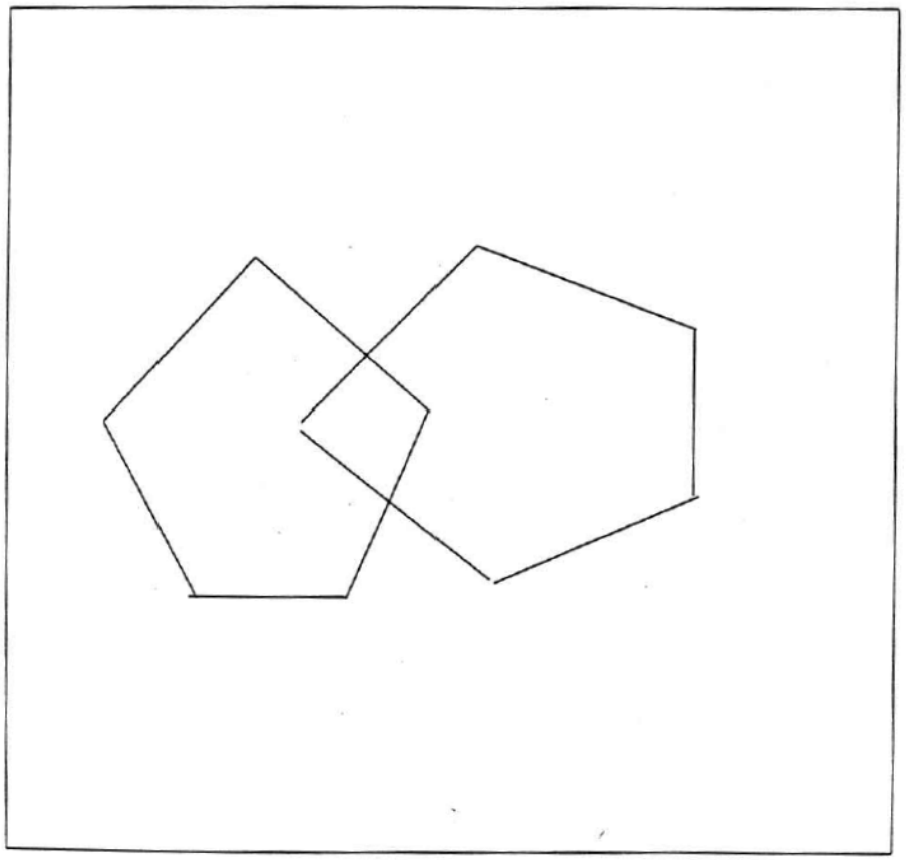
Total Score: _____ / 30

< 24 Considered Abnormal
< 18 Do Not Do Time-Trade Off
< 16 Do Not Do Geriatric Depression Scale

Version 1 (July 7, 2011)

Figure B.10: Contents of the Follestini Mini-Mental Test used in all three of the studies of this thesis (page 2 of 3).

CLOSE YOUR EYES



Page 3 of 3

Version 1 (July 7, 2011)

Figure B.11: Contents of the Follestini Mini-Mental Test used in all three of the studies of this thesis (page 3 of 3).

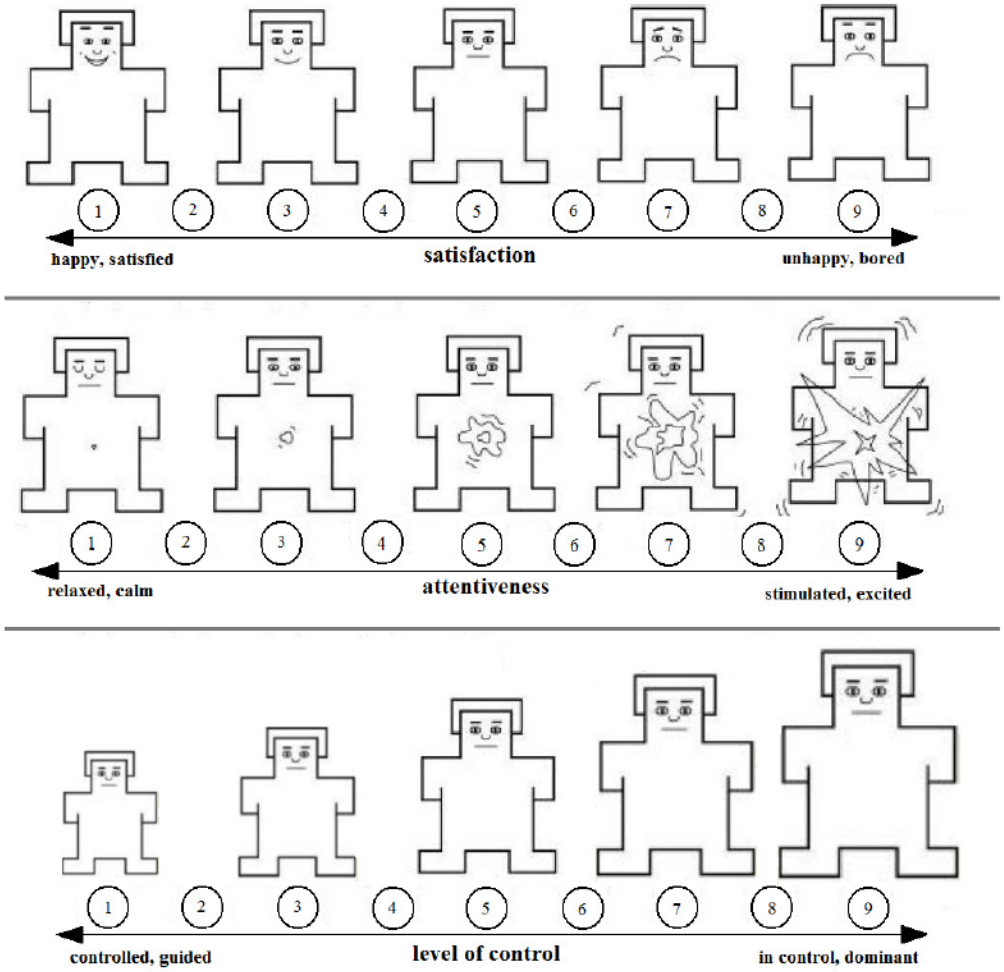
Post-trial Questionnaire

This questionnaire is designed to gauge your feelings about the experience of doing the session of tasks that you just completed. This questionnaire is a picture-oriented instrument called the Self-Assessment Manikin (SAM) to directly assess the associated satisfaction, attentiveness and level of control in response to/over an object or event. The questions relate to how much you liked the trials, how attentive you were and how in control you felt, each scale represented by a row of figures in the questionnaire. The data from this questionnaire will be used to assess the overall level of your engagement in the task. **You are not required to answer any questions you do not feel comfortable answering.**

For each of the three dimensions, please choose the one choice (out of the nine possible choices) that describes your feeling towards the task you just completed the best.

Version 2 (June 23, 2011)

Figure B.12: The Self Assessment Manikin Questionnaire used in the studies presented in Chapter 2 and Chapter 4 (page 1 of 2).



Version 2 (June 23, 2011)

Figure B.13: The Self Assessment Manikin Questionnaire used in the studies presented in Chapter 2 and Chapter 4 (page 2 of 2).

Subject #: Date:

Post-experiment Questionnaire

You are not required to answer any questions you do not feel comfortable answering

A. Demographics

1. What is your age? _____
2. What is your gender? Female / Male
3. What is your dominant hand? Right-handed / Left-handed
4. How familiar are you with working / interacting with a robot?

Not familiar	1	2	3	4	5	Very familiar

5. Have you ever worked / interacted with any robots? If yes, please describe your experience with the robot:

Version Date: August 8, 2012

Figure B. 14: The post-experiment questionnaire used in the study presented in Chapter 4 (page 1 of 4).

Subject #: Date:

B. Task Load

1. Mental Demand — How mentally demanding was the task?

Very Low	1	2	3	4	5	Very High

2. Physical Demand — How physically demanding was the task?

Very Low	1	2	3	4	5	Very High

3. Performance — How successful were you in accomplishing what you were asked to do?

Failure	1	2	3	4	5	Perfect

4. Effort — How hard did you have to work to accomplish your level of performance?

Very Low	1	2	3	4	5	Very High

5. Frustration — How insecure, discouraged, irritated, stressed, and annoyed were you?

Very Low	1	2	3	4	5	Very High

C. Describe the Robot

For each word below, please indicate how well it describes the ROBOT you just worked with:

	Describes very poorly				Describes very well
Independent	1	2	3	4	5
Helpful	1	2	3	4	5
Assertive	1	2	3	4	5
Competitive	1	2	3	4	5
Dominant	1	2	3	4	5
Forceful	1	2	3	4	5

Figure B.15: The post-experiment questionnaire used in the study presented in Chapter 4 (page 2 of 4).

Subject #: Date:

D. Describe the Interaction

For each word below, please indicate how well it describes your INTERACTION with the robot:

	Describes very poorly				Describes very well
Boring	1	2	3	4	5
Enjoyable	1	2	3	4	5
Engaging	1	2	3	4	5

1. In this experiment you finished five training sets of reaching tasks. If there was no time limit for the experiment, how long (how many more training sets) would you continue this robotic task? _____
2. I lost myself in this robotic experience.

Strongly Disagree	1	2	3	4	5	Strongly Agree

3. This robotic experience was fun.

Strongly Disagree	1	2	3	4	5	Strongly Agree

4. I felt involved in this robotic task.

Strongly Disagree	1	2	3	4	5	Strongly Agree

5. This robotic experience did not work out the way I had planned.

Strongly Disagree	1	2	3	4	5	Strongly Agree

6. I was so involved in this robotic experience that I lost track of time.

Strongly Disagree	1	2	3	4	5	Strongly Agree

7. I felt discouraged while interacting with the robot.

Strongly Disagree	1	2	3	4	5	Strongly Agree

Version Date: August 8, 2012

Figure B.16: The post-experiment questionnaire used in the study presented in Chapter 4 (page 3 of 4).

Subject #: Date:

8. I felt frustrated while interacting with the robot.

Strongly Disagree	1	2	3	4	5	Strongly Agree

9. I felt in control of this robotic experience.

Strongly Disagree	1	2	3	4	5	Strongly Agree

E. Impression and Attitude

1. Please rate your impression of the ROBOT on these scales:

Mechanical	1	2	3	4	5	Organic
Pleasant	1	2	3	4	5	Unpleasant
Intelligent	1	2	3	4	5	Unintelligent
Like	1	2	3	4	5	Dislike

2. I will feel uneasy if I am given a job where I have to use robots.

Strongly Disagree	1	2	3	4	5	Strongly Agree

3. I hate it when robots or artificial intelligence make judgments about things.

Strongly Disagree	1	2	3	4	5	Strongly Agree

4. I feel very nervous just standing in front of a robot.

Strongly Disagree	1	2	3	4	5	Strongly Agree

5. I feel that if I depend on robots too much, something bad might happen.

Strongly Disagree	1	2	3	4	5	Strongly Agree

Version Date: August 8, 2012

Figure B.17: The post-experiment questionnaire used in the study presented in Chapter 4 (page 4 of 4).

Appendix C - Cogging Compensation: Modeling, Identification, and Controller Design

The robotic manipulandum introduced in this thesis was used as a mean to investigate effects of distorted feedback on enhancing motor adaptation. During the pretests of the study presented in Chapter 2 (study I), participants complained about a disturbing vibration as they interacted with the manipulandum.

Albeit having excellent features such as high torque to current ratio and fast response, one of the motors shows rather high friction and cogging amplitudes as illustrated in Figure C.1. The vibration that subjects complained about was caused by high cogging torque of this motor which was discovered by moving the end effector in a way that only excited a movement in that motor.

Friction and cogging forces could bias the proposed studies of this thesis by disturbing participants. This appendix looks into modeling and identification of friction and cogging torques, followed by implementation of a friction compensation feature on the current controller of the robot.

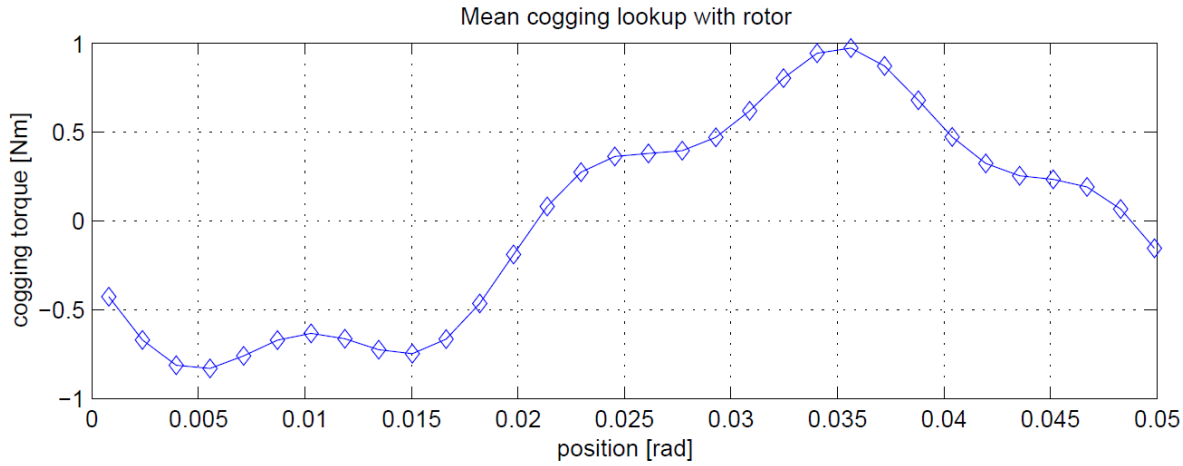


Figure C. 1: One of the motors shows a noticeable cogging torque.

C.1 Methodology and Implementation

Three main phases of this study are:

1. Proposing a model for friction and cogging torques in the motor,
2. Identification process to find numerical values for coefficients of the proposed model,

3. Implementation and tuning of a feedforward compensator based on the previous steps.

Atsma, designer of the robotic manipulandum, reviews dynamic equations of the manipulandum and suggests a linearised extended model for required torque to maintain a desired trajectory (position, velocity and acceleration) using Lagrangian method in [C1]. He enhances this model with friction and cogging models.

Atsma formulates dynamics of the manipulandum in a matrix format as:

$$\tau = H(q)\ddot{q} + C(q,\dot{q})\dot{q} \tag{C.1}$$

where τ is torque applied by motors, $q, \dot{q}, \ddot{q}$ represent position, velocity and acceleration, $H(q)$ is the inertia matrix and finally $C(q,\dot{q})$ has the coriolis and centrifugal components.

It is more convenient to represent the above equation in a linearised equation as multiplication of two matrices, one formed based on data from the encoders (current position) and the reference trajectory that the robot is to follow (inferred from the force operator applies on the end effector, measured by the ATI sensor), the other containing information on inertia of the links:

$$\tau = Y_r(q,\dot{q},\dot{q}_m,\ddot{q}_m)a_r = H(q)\ddot{q} + C(q,\dot{q})\dot{q} \tag{C.2}$$

This is implemented in the original impedance controller of the manipulandum (Figure C.2) [C2]. This controller outputs a toque command to each of the two motors based on the force operator is applying to the end effector (desired motion), predesigned perturbation criteria which was used in previous studies conducted by Atsma, and position of the end effector ("bounding box" box in the model ensures operator stays within the working area). The original controller [C2] does not compensate for friction, as it was assumed that friction and cogging are negligible.

Equation C.2 is valid only if friction is negligible. To account for effects of friction in actuator *i*, a simple friction model can be added to equation C.2 to compensate for static and viscous friction:

$$\tau_{fi} = b_{f\_viscous\_i}\dot{q}_i + \tau_{f\_static\_i} \tag{C.3}$$

where $b_{f\_viscous\_i}$ is the viscous friction in actuator *i* and $\tau_{f\_static\_i}$ is the static friction.

Assuming that the cogging profile is sinusoidal, the model for cogging torque is then simply the sum of a sine and a cosine component with unknown amplitude coefficients and cogging frequency, which will be added to equation (C.2) as well:

$$\tau_{cog1} = a_1 \cos(q_1/f_{cog}) + a_2 \sin(q_1/f_{cog}) \quad \text{(C.4)}$$

$$\tau_{cog2} = a_3 \cos(q_2/f_{cog}) + a_4 \sin(q_2/f_{cog}) \quad \text{(C.5)}$$

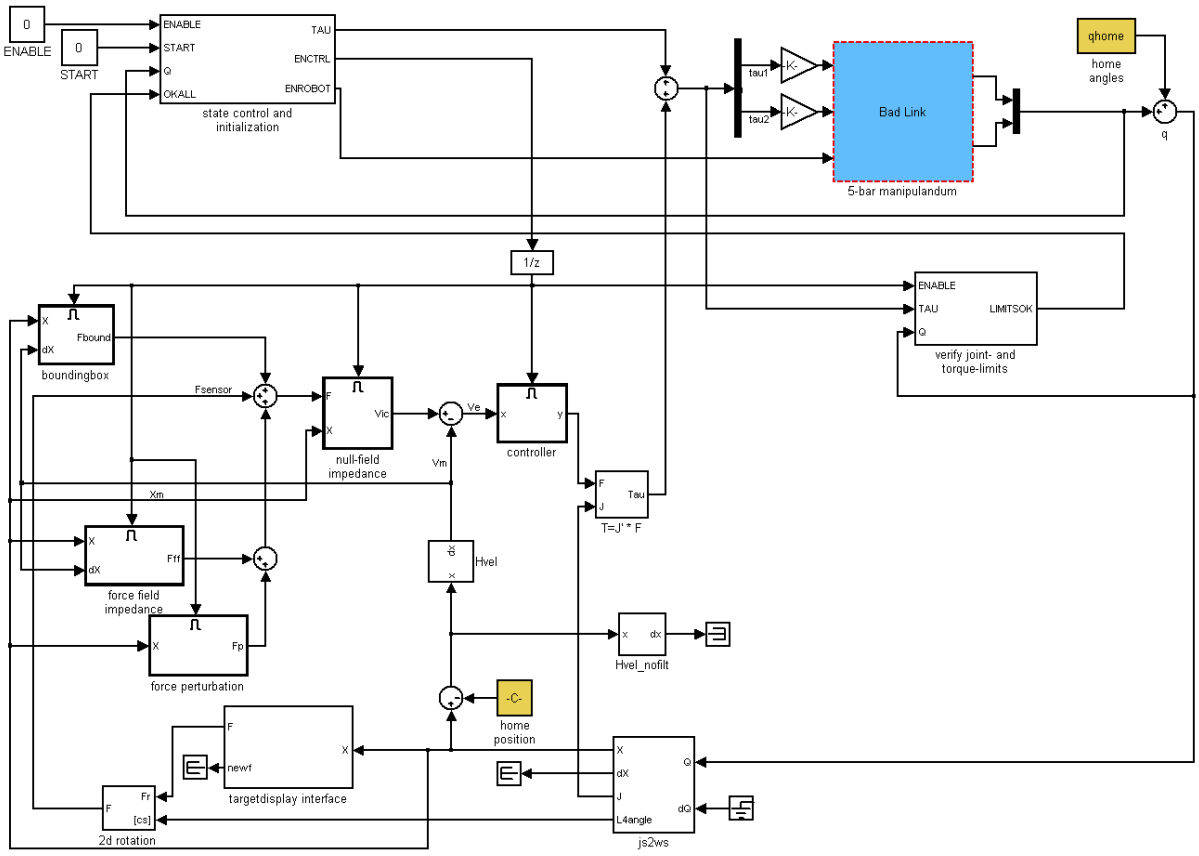


Figure C. 2: Simulink model of the original impedance controller of the manipulandum.

However, after going through identification phase, results (Figure C.1) suggested that it is better to model cogging as sum of two sinusoids with two different frequencies:

$$\tau_{cog1} = a_1 \cos(q_1/f_{cog1}) + a_2 \sin(q_1/f_{cog1}) + a_3 \cos(q_1/f_{cog2}) + a_4 \sin(q_1/f_{cog2}) \quad \text{(C.6)}$$

$$\tau_{cog2} = a_5 \cos(q_2/f_{cog1}) + a_6 \sin(q_2/f_{cog1}) + a_7 \cos(q_2/f_{cog2}) + a_8 \sin(q_2/f_{cog2}) \quad \text{(C.7)}$$

This model is linear and simple enough to be implemented. However, to be able to use feed-forward control to cancel out unwanted dynamic parameters, accurate knowledge of friction and cogging parameters are required.

To design a high performance CNC, Erkorkmaz and Altintas [C3] propose a method to estimate dynamics of the feed drives (system) and develop a friction model. In their paper, after modeling the linear dynamics of a feed drive and main sources of error, the physics of the friction phenomenon is briefly discussed and a model for friction is introduced.

To find the numerical values of the friction model, the inertia, Coulomb friction and viscous friction of each feed axis (previously referred to as dynamics of the feed) are estimated through an unbiased least squares technique. 'Unbiased' here means that the effects of noise and Coulomb friction were considered in discrete time state space equations used for the estimation process. Values obtained from biased and unbiased estimation are compared and unbiased estimation shows more consistency.

As our primary solution, we used unbiased least squares estimation to estimate friction and cogging coefficients of the manipulandum as modeled by Atsma. To do so, we will use the same formulation Erkorkmaz [C3] has used for static and viscous friction.

Figure C.3 shows the simplified linear dynamics of a typical motor. T_m is the torque generated by the commanded voltage and current running through the armature. In addition to this torque, the motor shaft is also subject to a disturbance torque containing effects of friction and the load. Based on this model, discrete state space equations can be written as in Equation C.8.

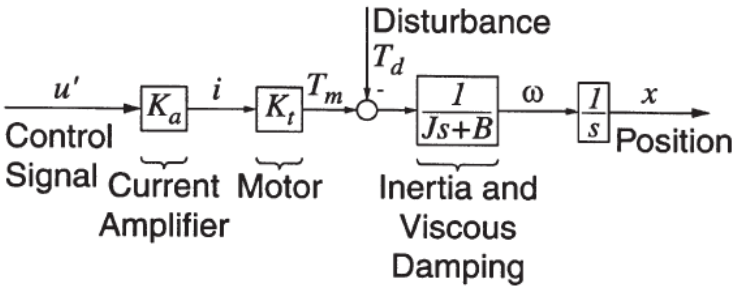


Figure C. 3: Linear dynamics of a motor

$$\omega(k+1) = p_{\omega d}\omega(k) + [K_{\omega d} \quad -K_{\omega d}]\begin{bmatrix} u'(k) \\ d(k) \end{bmatrix}, p_{\omega d} = e^{p_\omega T_s}, K_{\omega d} = \int_0^{T_s} e^{p_\omega \gamma}.d\gamma.K_\omega \quad \text{(C.8)}$$

where $d(k) = T_d(k)/(K_t K_a)$, $p_\omega = -B/J$ and $K_\omega = K_t K_a/J$. Erkorkmaz using exactly the same equation introduced his unbiased identification method. Assuming that the measurement of axis velocity $\omega_m(k)$ is corrupted with Gaussian noise $\tilde{\omega}(k)$ and the effective input signal $u'(k)$ is corrupted with uniformly distributed noise $\tilde{u}(k)$ as:

$$\begin{cases} \omega_m(k) = \omega(k) + \tilde{\omega}(k) \\ u'(k) = u(k) + \tilde{u}(k) \end{cases} \tag{C.9}$$

then Equation C.8 can be rewritten as:

$$\omega_m(k) = p_{\omega d}\omega(k) + K_{\omega d}u(k-1) - K_{\omega d}d(k-1) + \tilde{\omega}(k) - p_{\omega d}\tilde{\omega}(k) - 1 + \tilde{u}(k-1) \tag{C.10}$$

Of the bias sources, the disturbance term of $K_{\omega d}d(k-1)$ has the most severe impact. Assigning a constant value to $d(k)$ to model Coulomb friction we will get:

$$d(\omega(k)) = \begin{cases} 0 \; if \; \omega(k) = 0 \\ d^+ \; if \; \omega(k) > 0 \; (d^+ = \frac{T_{Coul}^+}{K_t K_a}) \\ d^- \; if \; \omega(k) < 0 \; (d^- = \frac{T_{Coul}^-}{K_t K_a}) \end{cases} \tag{C.11}$$

In the estimation process to distinguish between true non-zero values of velocity and noise in measurement a dead band just above the maximum level of noise is used instead of comparing the velocity to zero.

Introducing PV and NV as:

$$\begin{cases} PV: Positive\ Velocity = 1 \; if \; velocity \; is \; above \; positive \; deadband \\ NV: Negative\ Velocity = -1 \; if \; velocity \; is \; below \; negative \; deadband \end{cases} \tag{C.12}$$

Equation C.10 can be rewritten:

$$\omega_m(k) = \begin{bmatrix} \omega_m(k) & u(k-1) & -PV(\omega_m(k-1)) & -NV\big(\omega_m(k-1)\big) \end{bmatrix} \begin{bmatrix} p_{\omega d} \\ K_{\omega d} \\ K_{\omega d}d^+ \\ K_{\omega d}d^- \end{bmatrix} \tag{C.13}$$

Equation C.13 can be rewritten for N collected data samples as:

$$Y^0 = \emptyset^0 . \theta^0 + E^0 \tag{C.14}$$

in which Y^0 is the output vector, θ^0 is the parameter vector, $\emptyset^0$ is the matrix of regression and E^0 is the prediction error vector. Then the parameter vector estimate can be obtained as:

$$\hat{\theta}^0 = ((\emptyset^0)^T \emptyset^0)^{-1} (\emptyset^0)^T Y^0 \tag{C.15}$$

The first entry of $\hat{\theta}^0$ will be $\hat{p}_{\omega d}$ and the second $\widehat{K}_{\omega d}$. d^+ and d^- are then obtained by dividing the third and fourth entries by $\widehat{K}_{\omega d}$ respectively. And finally Coulomb friction can be derived from these two values by multiplying them with $K_t K_a$. From $\hat{p}_{\omega d}$ and $\widehat{K}_{\omega d}$ values for B and J can be obtained. For more details see [C3].

Identification tests on the cogging motor were conducted. These involved commanding a series of step inputs and analyzing the data to find values for J, B and Coulomb friction that can then be used as coefficients of Equation C.3.

The discrete time state space equation of the motor can be written as:

$$\begin{bmatrix} x(k+1) \\ \omega(k+1) \end{bmatrix} = A_d \begin{bmatrix} x(k) \\ \omega(k) \end{bmatrix} + \begin{bmatrix} B_d & -B_d \end{bmatrix} \begin{bmatrix} u'(k) \\ d(k) \end{bmatrix} \tag{C.16}$$

Considering the noisy measurements, Equation C.16 can be rewritten as:

$$\begin{cases} \begin{bmatrix} x(k+1) \\ \omega(k+1) \\ d(k+1) \end{bmatrix} = A \begin{bmatrix} x(k) \\ \omega(k) \\ d(k) \end{bmatrix} + B[u(k)] + W \begin{bmatrix} \tilde{u}(k) \\ \omega_d(k) \end{bmatrix} \\ \begin{bmatrix} x_m(k) \\ \omega_m(k) \end{bmatrix} = C \begin{bmatrix} x(k) \\ \omega(k) \\ d(k) \end{bmatrix} + V \begin{bmatrix} \tilde{x}(k) \\ \tilde{\omega}(k) \end{bmatrix} \end{cases} \tag{C.17}$$

Where:

$$A = \begin{bmatrix} A_d & & -B_d \\ 0 & 0 & 1 \end{bmatrix}, B = \begin{bmatrix} B_d \\ 0 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, W = \begin{bmatrix} B_d & 0 \\ & 0 \\ 0 & 1 \end{bmatrix}, V = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{C.18}$$

Values in Equation C.18 can be easily calculated from the results of the identification test; details are omitted here. The state space model of Equation C.17 can be used to estimate motor position *x*, velocity ω and disturbance *d* in terms of a Kalman Filter as:

$$\begin{bmatrix}\tilde{x}(k+1)\\ \tilde{\omega}(k+1)\\ \tilde{d}(k+1)\end{bmatrix} = (I - K_{obs}C)A\begin{bmatrix}x(k)\\ \omega(k)\\ d(k)\end{bmatrix} + (I - K_{obs}C)B[u(k)] + K_{obs}\begin{bmatrix}x_m(k+1)\\ \omega_m(k+1)\end{bmatrix} \quad \text{(C.19)}$$

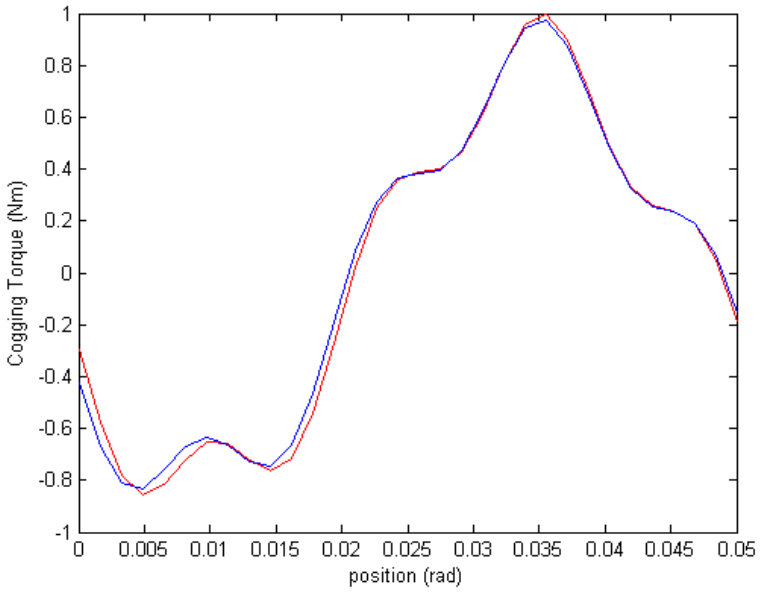


Figure C. 4: Comparison between actual data (blue) and fitted model (red), a cogging cycle is 2π/124 rads.

Cogging torque (position dependent) can be considered as the difference between an estimated disturbance value from a Kalman filter and a fitted value (Coulomb friction) from the unbiased estimation. Equation C.6 was fitted to resulting data from this operation to find numerical values for the coefficients of the cogging model. The cogging frequencies used are 124 and 4*124 (spatial frequencies), which are the number of poles per revolution and 4 teeth per pole. Therefore, a cogging cycle is 2π/124 (rad), almost 3 degrees. Figure C.4compares the data from the estimation of cogging and the fitted model. Given the small scale of each cogging cycle, the differences between these two values are negligible.

Results of the identification process for the cogging motor are as follows:

$K_t = 146.73$

$K_a = 1$

$J = 0.03762\ kgm^2$

$B = 0.8539\ Nms$

$T_{Coul}^{+} = 1.6515\ Nm$

$T_{Coul}^{-} = -1.7896\ Nm$

$a_1 = -0.3401\ Nm$

$a_2 = -0.7553\ Nm$

$a_3 = 0.0454\ Nm$

$a_4 = -0.1702\ Nm$

Having the numerical values for the coefficients of the friction and cogging models, a compensation box was added to the Simulink model of the controller previously shown in Figure C.2. Inputs to this compensator are motor position and motor velocity. Output is a feedforward (compensating) torque that is added to the commanded torque (see Figure C.5). Figures C.6 and C.7 show a breakdown of the compensation box.

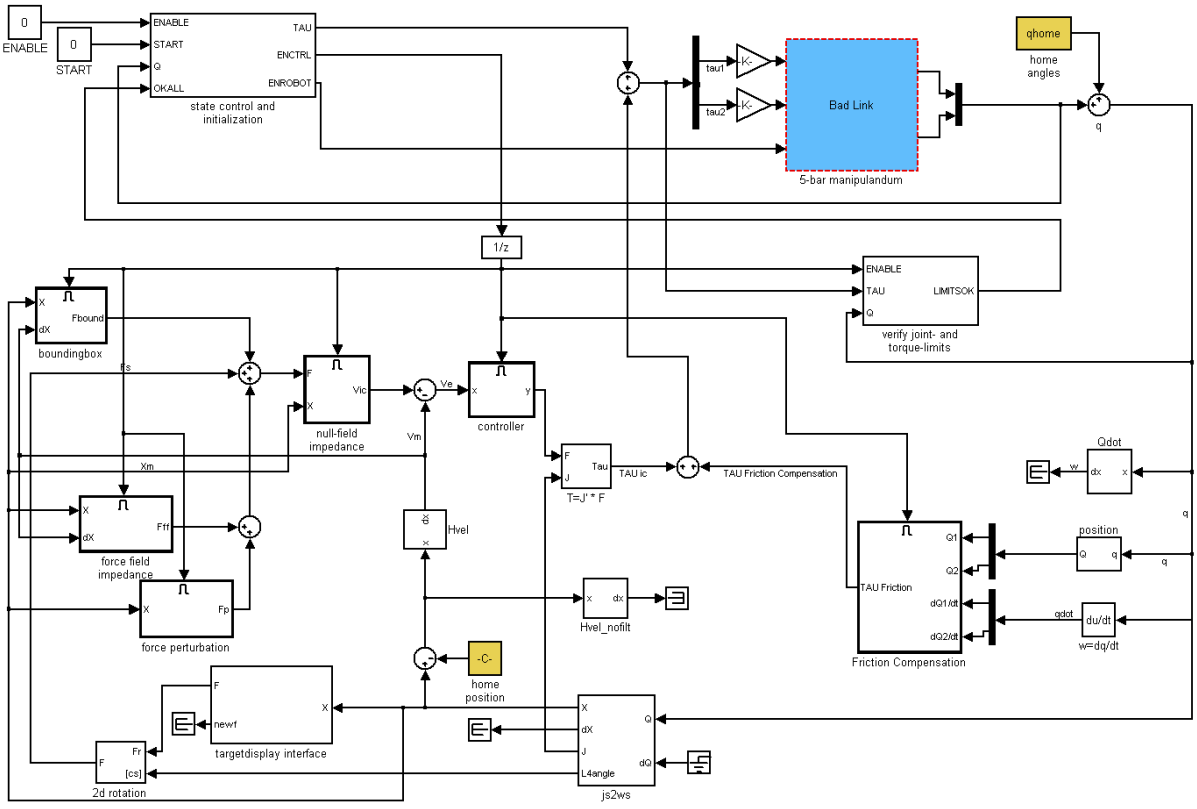


Figure C. 5: A friction compensation box was added to the Simulink model based on findings of identification process.

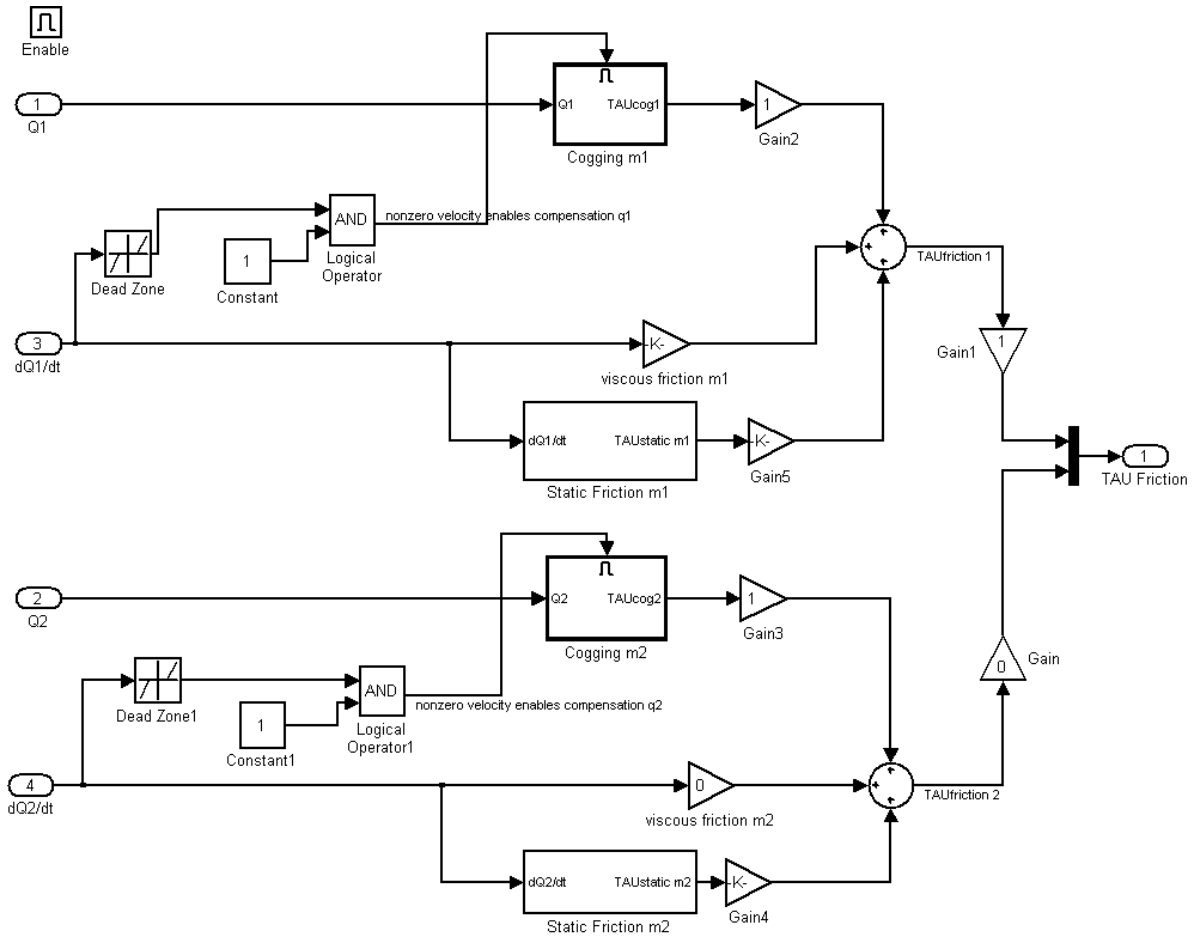


Figure C. 6: Details of the cogging compensator box. Based on position and velocity of each motor compensator box sends out a commanded torque to compensate for cogging, viscous friction and Coulomb friction.

Figure C.6shows components within the compensator box. Inputs are the position and velocity of the motor. Based on these inputs and using the models introduced above, the compensator box sends out a commanded torque as an output to compensate for cogging, viscous friction and Coulomb friction. The same identification process was done for the non-cogging motor ($q_2$and dq_2 in Figure C.6) but results are not presented here. As the current performance of this motor is satisfactory, no compensation is done for this motor (note a gain of zero is being applied on "TAUfriction 2" signal).

Figure C.7shows components within the "Cogging m1" box in Figure C.6. Box is enabled only if the motor is moving. Input is the position (q_1) of the cogging motor. q_1 is scaled down to cogging cycles and a phase shift is added to fine tune the compensator to sync the compensator's cycle with the actual cogging cycle, results of which are presented in the next section.

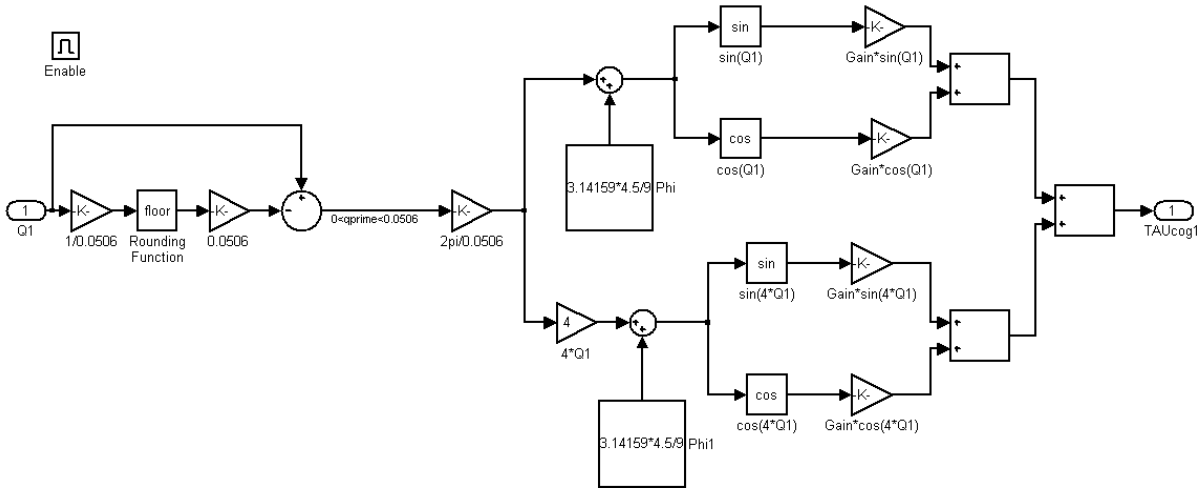


Figure C. 7: Details of "Cogging m1" box of Figure C.6.

C.2 Results and Comparison

To look at the changes in performance of the system, the operator was asked to move the representation of the end effector on the system's monitor (depicted by a cursor) from the midpoint to a target point and back to the midpoint. The target point was chosen in such a way that moving to it only required rotation of the cogging motor. In theory, measurement of the applied force to the end effector to do this task must have a bell shape [C4]. But due to vibration, especially in the case where cogging is an active disturbance, this applied force will be jittery, causing the standard deviation of the data to grow larger.

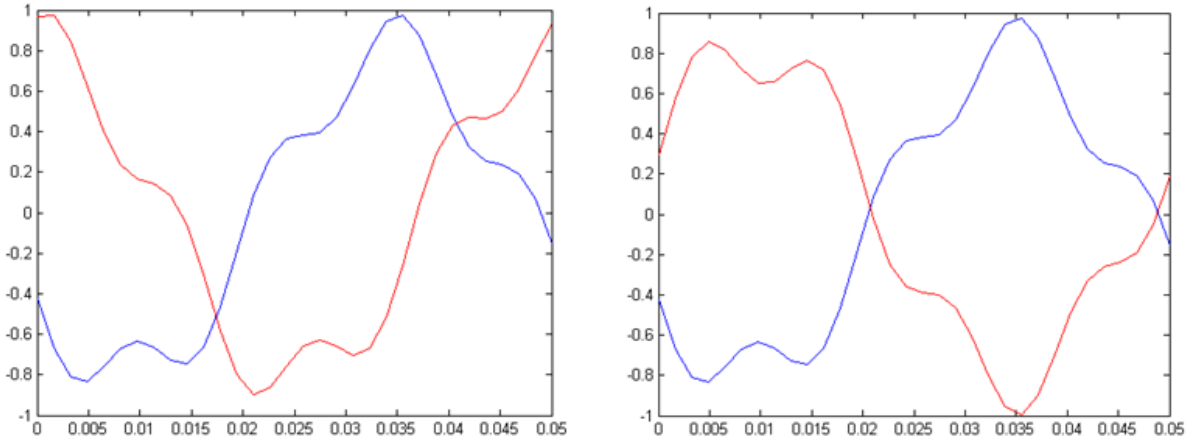


Figure C. 8: Phase shift matters! Red curve shows calculated cogging torque and blue curve is actual cogging torque of motor. These two have to be exactly 180 degrees out of phase in order for the cogging torque to cancel out (right plot).

Although the model of the cogging friction is accurate, simply feeding in motor position to this model does not work. As Figure C.8shows, the blue curve is the actual cogging torque from the motor, while the red curve is the calculated cogging torque based on encoder position. The

curves are the same but the red curve has a phase shift compared to the blue one. This phase shift can play a significant role in cancelling out the cogging torque. Different phase shifts (5° to 180°with a step of 5°) have been implemented in the "Cogging m1" box (Figure C.7) to find the case yielding the smallest standard deviation in data of force measurement.

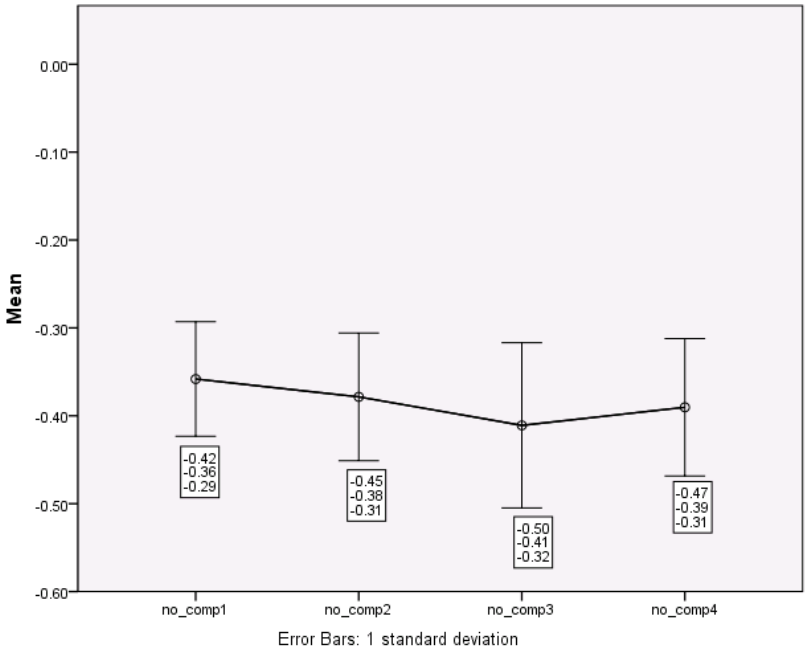


Figure C. 9: Results of running the motor without compensation, average standard deviation is 0.8 N.

As indicated by Figure C.9, initially and without any friction compensation, an average standard deviation of 0.8N was observed (the experiment was run four times). By adding compensation, the best result was obtained using a phase shift of 80° (Figure C.10). Also, the mean values of exerted force in all of the experiments were in the same range (ideally they should be equal), as the operator was asked to do the task with a consistent pace throughout the experiment.

After the tuning process, a compensator with an 80° phase shift was implemented in the manipulandum's controller. Interacting with the manipulandum and moving it around was reported by the test subjects to be smoother than without the compensator.

To further develop this work, it would be beneficial to model cogging parameters as both position and velocity dependent. Also, while running the system, it often happened that the system became sluggish and acted as is there was no active cogging compensation. This would disappear upon restarting the NT machine. This known "bug" is absent in newer versions of MATLAB, so a software upgrade is recommended.

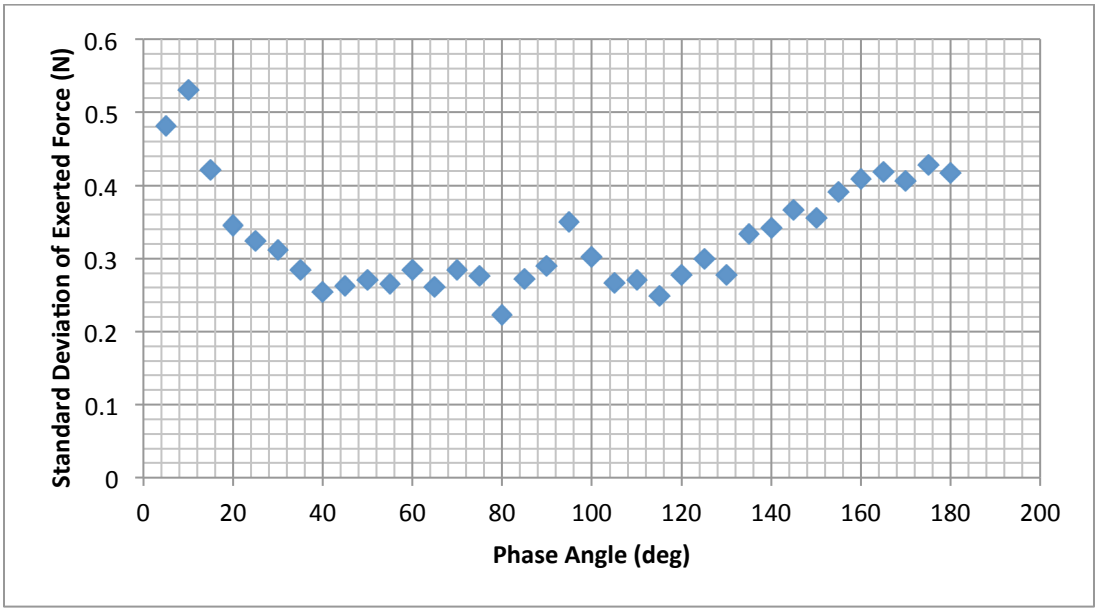


Figure C. 10: Standard deviation of exerted force for different phase angles. 80 degrees shows the smallest S.D.

References

[C1] W.J. Atsma, "Inference of central nervous system input and its complexity for interactive arm movement," Ph.D. Dissertation, Department of Mechanical Engineering, The University of British Coulombia, 2006.

[C2] W.J. Atsma, "Lagrangian Dynamics Equations and a Linearised Extended Model."

[C3] Y. Erkorkmaz, K; Altintas, "High speed CNC system design. Part II: modeling and identification of feed drives," International Journal of Machine Tools and Manufacture, Vol. 41, 2001, pp. 1487-1509.

[C4] T. Flash, N. Hogan, "The Coordination of Arm Movements: An Experimentally Confirmed Mathematical Model", The Journal of Neurosciences, Vol. 5, No. 7, 1985, pp. 1688-1703.