**Title**
Single-Site Surface Electromyography for Human-Machine Interfaces



**Author**
O'Meara, Sarah

**Publication Date**
2021

Peer reviewed|Thesis/dissertation

Single-Site Surface Electromyography for Human-Machine Interfaces

By

SARAH MARIA MENDES O'MEARA

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Mechanical and Aerospace Engineering

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Stephen K. Robinson, Co-Chair

Sanjay S. Joshi, Co-Chair

Bahram Ravani

Committee in Charge

2021

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

**Single-Site Surface Electromyography for Human-Machine Interfaces**

The aim of my dissertation was to investigate the command and control of machines by humans. There have been many control methods to interface with machines, and this dissertation focused on surface electromyography (sEMG). Surface EMG is the measurement of electrical signals produced by muscle(s) and measured at the skin's surface. Several EMG techniques rely on multiple sensors and machine learning for multi-DOF (degree of freedom) or multi-command control and are therefore more complex and may be sensitive to signal degradation. The first goal was to develop a robust, single-site sEMG control methodology that could communicate more than one command from a single muscle site or EMG sensor. A computer-based cursor-to-target task assessed the performance of the single-site sEMG for two levels of control: auto-rotate (automatic cursor rotation and user-controlled cursor forward motion) and manual rotate (user-controlled cursor rotation and forward motion). The experiment demonstrated that the auto-rotate method led to better throughput, or the rate of selecting targets adjusted for difficulty, but not significantly better path efficiency. However, the subjects were able to learn the manual-rotate method where a single-site sEMG control method communicated two commands. The manual rotate method appeared viable for expansion from two to four commands while maintaining a single sensor, but it was unclear how to best train subjects to use the system. The subsequent experiment investigated the effects of different training methodologies on performance, cognitive workload, and trust. The augmented feedback training techniques led to early and sustained performance gains with lower cognitive workload and higher trust. Even with extensive training to learn the commands, subjects did not perfectly perform the four sEMG commands. A subsequent analysis revealed that adjusting the sEMG command parameters may help personalize the command system to the individuals and improve command performance. The knowledge gained from these studies culminated in a three-handed coordination pilot study with a sEMG-controlled, collaborative robot serving as the third hand. Subjects learned how to improve their coordination of the three "hands." The pilot study also served to inform future experiments that will continue to integrate the human with a collaborative robot for more complex tasks.

# ACKNOWLEDGEMENTS

To those who encouraged and challenged me—thank you for the experiences and growth opportunities. Thank you to the subjects for volunteering to participate in these research studies.

This research was conducted at the University of California, Davis in the Human, Robotics, Vehicle Integration & Performance Laboratory and the Robotics, Autonomous Systems, and Controls Laboratory. Funding was provided by the National Science Foundation (1208186, 1934792), NASA's Space Technology Research Grant Program, the Link Foundation, and the Zonta International Foundation. The contents of this dissertation have been previously published in [1]–[4].

# Chapter 1

# Introduction

## 1.1 Body-Machine Interfaces for Aerospace Applications

Human-Machine Interfaces (HMIs) traditionally use input devices such as a computer mouse, joystick, and steering wheel. A more recent subset of HMIs leverages signals derived from the human body to communicate intent and has been termed as Body-Machine Interfaces (BoMIs) [5]. These BoMIs can be passive, where the body signals are used as a state indicator to monitor or trigger a response to the human, or active, where the human directly controls the input to determine the device output. Furthermore, BoMIs are characterized by the body signal they use, such as brain signals, gestures, or muscular signals, and whether the measurement is invasive (i.e., penetrates the body) or noninvasive. BoMIs that use brain signals may be termed as Brain-Machine Interfaces (BMIs) or Brain-Computer Interfaces (BCIs), where the latter is a BMI specifically designed for computer-based applications. The present discussion is focused on noninvasive techniques for aviation and space applications.

An increased interest in BoMIs for use in aviation has led to the development of several passive and active systems. Passive systems have measured brain signals to predict the occurrence of inattentional deafness to auditory alarms [6], cognitive fatigue [7], monitor cognitive load [8], and track vigilance [9]. For example, Di Flumeri et al. [9] demonstrated improved vigilance when a BCI system detected changes in subjects' vigilance and automatically adjusted the automation level to increase/decrease manual engagement for an air traffic control task. Electromyography (EMG), the recording of electrical muscular signals, has also been used for passive BoMIs to warn of gravity-induced loss of consciousness [10], [11], monitor muscle fatigue [12], [13], and differentiate pilot skill level [14]. The prior exemples are for *passive* EMG BoMIs; however, *active* EMG BoMIs are less prevalent in aviation. The work by Jorgensen et al. [15] demonstrated the successful implementation of an EMG BoMI to control the pitch and bank rates for an aircraft landing in simulation. In contrast, there has been an increased ubiquity in the use of body gestures

for active BoMIs related to drone control (see Refs. [16]–[19] for some recent examples). One study observed significantly better performance with torso gestures than a traditional HMI, a joystick [18], lending to the feasibility and potential benefits of these nontraditional interfaces.

In 2004, the European Space Agency's Advanced Concept Team identified noninvasive BMIs as a research area with potential space applications [20]. Broschart et al. [21] proposed incorporation of BMIs for teleoperations, assistive robots, and to increase astronaut productivity, but recognized that the unique physical environment coupled with physiological changes experienced by astronauts would pose implementation challenges. Some initial steps have occurred, such as a BCI mouse for online control of a simplified, simulated spacecraft in yaw and pitch [22]. However, the findings from this study were limited due to low subject number and the simplified simulation but indicate an interesting application. Another preliminary study with limited subjects used a BMI during parabolic flights, which simulate microgravity, and found that the brain signals measured by EEG (electroencephalography) remained stable enough to achieve an average classification accuracy of 73.1% [23]. In a review of BMIs for space applications, Coffey et al. [24] argued that the implementation of noninvasive, active BMIs was unlikely due to gaps in information transfer rates, accuracy, and intuitiveness. In contrast, passive BMIs were 1) less operationally demanding because they do not require active engagement or control, 2) had established reliability for monitoring aspects like attention, workload, and task engagement, and 3) the derived human state information could be fed into adaptable, autonomous systems [24]. BMIs may be more challenging to develop for space applications because of increased hardware complexity and relatively low signal-to-noise ratio (SNR). EEG signals measured from the surface of the scalp are attenuated by the structures between the brain and electrodes. Additionally, EEG signals are orders of magnitude smaller than other body signals, like those from muscles. Therefore, other BoMIs may be more feasible.

The use of muscular signals in *active* BoMIs has not been well-studied for aviation and space applications, but BoMIs that use EMG have multiple advantages. When EMG is measured from the skin's surface, it is noninvasive and referred to as sEMG. A sEMG system can be placed above any muscle location, can be designed to use a single sensor, and has a relatively high SNR. The sEMG signal quality

and reproducibility are affected by body pose [25], muscle fatigue [26], as well as sensor shifts [27]. These changes in the sEMG signal are also known as signal non-stationaries. Mitigating these issues is an active area of research (see review by Kyranou et al. [28]). Furthermore, EMG control has already been applied to prosthetics (see reviews by [29], [30]), as well as robotics (for examples, see [31]–[34]) which have potential aerospace applications (i.e. teleoperations [35]). In general, BoMIs provide a control modality that can be used in addition to or in replacement of traditional HMIs. Users are not constrained to use an input device with their hands (e.g. a computer mouse) and performance may improve with a BoMI (as seen in [18]).

### 1.2 Myoelectric Control

Myoelectric control, or the use of EMG to communicate commands to a device, has been used in upper limb prosthetics for more than fifty years, and commercially available prosthetics have been almost exclusively based on mode-switching and proportional control [29], [36]. In mode-switching the user may use one muscle group to turn on/off an action (e.g., close prosthetic hand) and another for a different action, or mode (e.g., open prosthetic hand) [29]. Proportional control uses EMG amplitudes as a relative input to the device (i.e., a larger EMG amplitude results in a higher joint velocity of the prosthetic). This type of control has been reliable and robust, which are priorities for patients and supersede the desire for the often less reliable multifunction control [29], but typically have a limited number of commands. In academia, researchers have focused on two strategies for multifunction control to increase the number of commands: pattern recognition for sequential commands and regression control for simultaneous commands [36] (refer to [30] for a comparison of features and methods). Pattern recognition has often been used to classify different gestures in able-bodied subjects to characterize performance, which can be measured by classification accuracy, number of gestures, and number of sensors. For example, one system achieved greater than 90% classification accuracy for six gestures using four sensors [37], whereas another classified four gestures with four sensors at an accuracy of 80% [38]. Both systems used Support Vector Machines

[39], a relatively simple classifier in terms of computational cost. Other systems used more complex classifiers and/or increase the number of sensors to train on larger datasets, however factors like computation cost should be considered in application (see discussion by [40]).

Single-site sEMG provides an alternative to increasing the system complexity with more sensors and complex classifiers. Our lab at the University of California (UC) Davis, the Robotics, Autonomous Systems, and Controls Laboratory (RASCAL), has been active in the development of single-site sEMG systems. The first system, a two-band technique, used the linear combination of the normalized power in two frequency bands (80-100 Hz, 130-150 Hz) to calculate the cursor x and y position on a screen [41]. The sEMG signal was measured from the extensor pollicis longus in the forearm. This initial case study demonstrated that the two-band technique was feasible; subjects were able to learn to modulate the signal power to move the cursor on the screen. In a subsequent case study, the technique was tested on a spinal cord injury patient using the auricularis posterior muscle located behind the ear, [42]. The patient was able to use the single-site sEMG control interface to move a cursor to select on-screen commands to navigate a robot through an obstacle course. Weisz et al. [43] extended this concept to a grasping task, where a spinal cord injury patient selected on-screen commands to select a target object, grasp, and then refine the grasp. Perez-Madonado et al. [44] also demonstrated the feasibility of this system using an adjacent muscle, the auricularis superior, in four able-bodied subjects. A larger study with twelve subjects was conducted to assess the learning rate during a cursor-to-target task using sEMG from the extensor pollicis longus, and demonstrated increased success rate and decreased completion time over the course of the multi-session study [45]. As an alternative to extracting multiple commands from a single sensor, the overall system design can be altered such that less commands are needed. Weisz et al. [46] designed to system to cycle through commands that could be selected once the cursor left a rest area. Most of the prior work in single-site sEMG used the two-band technique to derive two commands from one sensor. It was not clear if this technique could be expanded for more commands from the single sensors. The work presented in this dissertation explores an alternative single-site sEMG system that can accommodate more than two commands from one sensor.

## 1.3 Content Overview

The underlying motivation for this dissertation research was to explore the implementation of a single-site sEMG control method, where the learning and training burden was largely allocated to the human. Humans have the ability to learn complex skills and are able to quickly adapt. As a result, the myoelectric control design can be simplified and made more robust to signal non-stationaries. Although advanced machine learning can implement adaptive features or strategies to deal with signal non-stationaries, it comes at the cost of increased system complexity and reduced robustness. It is worthwhile to consider how to best use the human's innate abilities. From a human-machine integration perspective, proper allocation when designing roles would lead to better overall system performance. As discussed in the previous sections, there is an interest in alternative HMIs and the key to future integration and acceptance of BoMIs for aerospace applications lies in designing a high performance system.

This dissertation describes four studies concerning single-site sEMG using three muscle sites and between two and four commands. *Chapter 2* presents the initial development of this single-site sEMG system that is based on serial patterns of muscle activation. In *Chapter 3*, the system is expanded from the two commands in *Chapter 2* to four commands. Different training methodologies were assessed, as well as the impact on performance, cognitive workload, and trust. *Chapter 4* investigates the possibility of adding more complexity to the system through adaptable features and classifiers to the same system from the previous chapter. Finally, *Chapter 5* integrates robotic hardware with the sEMG system developed in previous chapters to assess coordination during a task between the hands and a lower limb using myoelectric control. The studies are summarized in the last chapter.

# Chapter 2

# Development of a Single-Site Surface Electromyography Control Method

## Preface

This chapter describes my first myoelectric study, where I extended our previous work in RASCAL on single-site surface electromyography methods for cursor control. I wanted to investigate the ability of subjects to learn a less intuitive cursor control method as opposed to a more intuitive one. This approach was of interest because a less intuitive cursor control method also meant that the subjects had the ability to make corrections, whereas corrections can be more difficult to make in intuitive systems reliant on machine learning. The cursor control method I developed had the potential to accommodate an increased number of commands without increasing the number of sensors. The cursor control task provided a test paradigm with standard, widely used metrics that focused on the subject's ability to learn a new myoelectric control scheme without the added infrastructure and complexity of a robotic testbed. The funding support for this experiment came from a grant focused on myoelectric control systems for patients with high-spinal cord injuries, which influenced the muscle site selection. The contents of this chapter have been previously published in [1].

## 2.1 Introduction

Developing single-site sEMG systems are not only of interest for downstream robotic applications but could benefit users with limited functionality who rely on Human-Computer Interfaces (HCIs). In general, HCIs bridge the gap between users and the computer. Cursor control is a popular HCI method which allows users to select commands on a screen. These commands may be instructions to an advanced prosthesis or onscreen letter selection for a speller application. Individuals with the most severe spinal cord injuries (C1-C4) often retain control of some head and neck muscles, and therefore several HCIs in the literature use head and neck muscle EMG as the input source [47], [48]. Typically, both the user-condition and the end-

goal application dictate the number and locations of EMG sensors. Users with severe spinal cord injuries may benefit from single-site sEMG systems due to reduced intrusiveness.

Previous sEMG HCIs have used several facial and neck muscles for user commands. Cler and Stepp capitalized on head muscle capability and demonstrated high efficiency using sEMG with a speller task application controlled by five facial muscles [49]. Each muscle functioned as a movement direction with the last one serving as a target selector or "clicker." Similarly, Huang et al. used facial gestures measured from four muscle sites to select between four directional and three selector commands [50]. Others, like Williams & Kirsch and Choi et al. developed a two degree of freedom (DOF) cursor control HCIs using EMG signals recorded from four head and neck muscles and forearm muscles, respectively [51], [52]. Williams & Kirsch utilized dwell time to select targets which effectively eliminated the need for a selection confirmation command. Similarly, Choi et al. were also able to reduce input commands by implementing a clockwise command and a counterclockwise command to eliminate the need for a downward command [52]. However, the exact muscle utilization would not be possible for many individuals with high-level spinal cord injuries due to the number and location of the muscle sites chosen. In this study, we emphasized the use of as few sensors as possible. Limiting the number of sensors makes the HCI less intrusive and could be applicable to greater numbers of people who have fewer functional muscle sites. The methods presented in this chapter only use a single muscle on the head.

The objective of this experiment was to develop and evaluate two distinct single-site, noninvasive EMG-modulated HCIs aimed at cursor control. It was also an opportunity to assess the viability of a less intuitive cursor control method. A Fitts's law-based task with a cursor-to-target paradigm was used to evaluate and compare performance between the two control methods, both of which enabled the cursor to move in 2-DOFs. The two methods differed in the level of user control, where one method allowed direct manipulation in 1-DOF and the other in 2-DOFs. For the first control method (the "auto-rotate" method), the cursor continuously rotates until the user chooses a favorable direction and then contracts the temporalis muscle above a specified threshold to send the cursor forward in the selected direction (see Figure 1 for electrode placement). For the second interface method (the "manual rotate" method), the user can choose

to rotate or move the cursor forward by using a series of muscle contractions of various length to indicate the command (similar to Morse code [53]). Target selection in both interfaces relied on dwell time to eliminate the need for an additional command.

We hypothesized that of the two control methods, the auto-rotate method would enable users to hit targets faster with lower workload. We further hypothesized that the manual rotate method would have more precision and perception of control over cursor behavior.

Fourteen able-bodied subjects with no prior EMG experience performed Fitts's law-based tasks for both control methods. Subjects also completed the National Aeronautics and Space Administration (NASA) Task Load Index (TLX) to assess cognitive workload, as well as surveys for demographics and subjective assessment of the control methods. The results suggest that a viable cursor control method can be achieved with only a single muscle site. The auto-rotate method is favorable with regards to performance and workload. However, subjects indicated similar control authority between the auto-rotate and manual rotate methods.

## 2.2 Methods

### *2.2.1 Experimental Design and Setup*

The experimental protocol was approved by the Institutional Review Board at UC Davis (#1281106); all subjects reviewed the protocol and consented prior to the experiment. Fourteen able-bodied subjects with no prior EMG control experience participated: six males/eight females with a combined average age of 26.64 ± 6.52 years ($\mu \pm \sigma$). All electrodes were placed on the dominant hand side with thirteen subjects having right-hand dominance.

Two disposable Ag/AgCl center snap electrodes (ConMed 1620) spaced approximately 2.5 cm apart were placed on the temporalis muscle and a gold disc electrode with conductive paste (Weaver and Company 10-20) was affixed to the earlobe as a reference (see Figure 1). The electrodes were connected to a Motion Lab Systems Y03 amplifier (x300 gain) powered by a custom power supply board and connected to a 16-bit Measurement Computing USB-1608G data acquisition unit following a protocol outlined in [54].

The signal was sampled at 4096 HZ in 256-sample windows and bandpass filtered with a fourth order digital Butterworth filter at 10 Hz and 500 Hz. The rms (root mean square) value for each window was calculated and normalized by a manually set calibration value to produce an updated signal input value, $x$, at 16 Hz. The result was then put into a moving average filter of length 8 to yield the fully processed signal input, $\bar{x}$ (used in Equation 1).

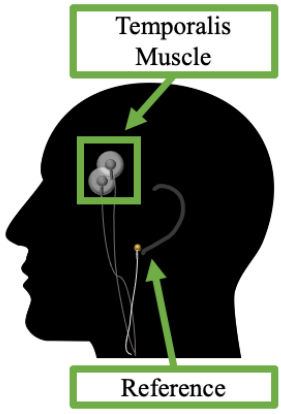



*Figure 1.* Electrode placement on temporalis muscle (signal source) and earlobe (reference).

Subjects were assigned to Group 1 or Group 2, which dictated the order of the cursor control methods presented, where Group 1 evaluated the auto-rotate method followed by the manual rotate method. In the auto-rotate method, the cursor automatically rotated at a constant speed and the user controlled the forward motion. The subjects controlled both the cursor rotation and forward motion in the manual rotate method. The experimental flow for subjects was as follows: consent form, a pre-session survey to collect demographics, experiment introduction, electrode placement, signal verification, manual calibration, method 1 instructions, method 1 training (8 trials), method 1 testing (80 trials), NASA-TLX, method 2 instructions, method 2 training (8 trials), method 2 testing (80 trials), NASA-TLX, and post-session survey for control method evaluation. A trial was defined as a single cursor-to-target task and testing was divided into blocks of 10 trials with two sessions (40 trials each). Subjects received mandatory breaks throughout training and testing to mitigate muscle fatigue. Please see *Appendix A* for the diagram of the experiment protocol.

### *2.2.2 Fitts's Law Task*

The overall user interface shown in Figure 2, included a task interface where the cursor could operate on the left-hand side, and a feedback interface on the right-hand side containing information about the sEMG signal and performance (throughput will be discussed in *2.2.4 Analysis*). The cursor interface dimensions were normalized to have horizontal and vertical bounds of [-1,1] on a square, right-hand Cartesian coordinate system. Target distances and widths are given in these normalized dimensions. At maximum velocity, the cursor could move from the center to the edge in 2 s.

Subjects used each control method for training and testing in the order determined by their Group. At the start of a trial, a target was presented at a randomized position and the cursor was positioned at the origin of the task interface with a randomized starting angle. There were five possible target distances (0.05, 0.1, 0.2, 0.4, 0.8) that produced a range of index of difficulties (ID) from 1.00 to 4.09 bits, as calculated by the Shannon formulation [55]. To complete the trial, the cursor had to dwell on the target for one second. There was a trial time limit of two minutes. For each control method during testing, the subject completed 80 trials (5 target distances × 8 target angles × two repetitions).

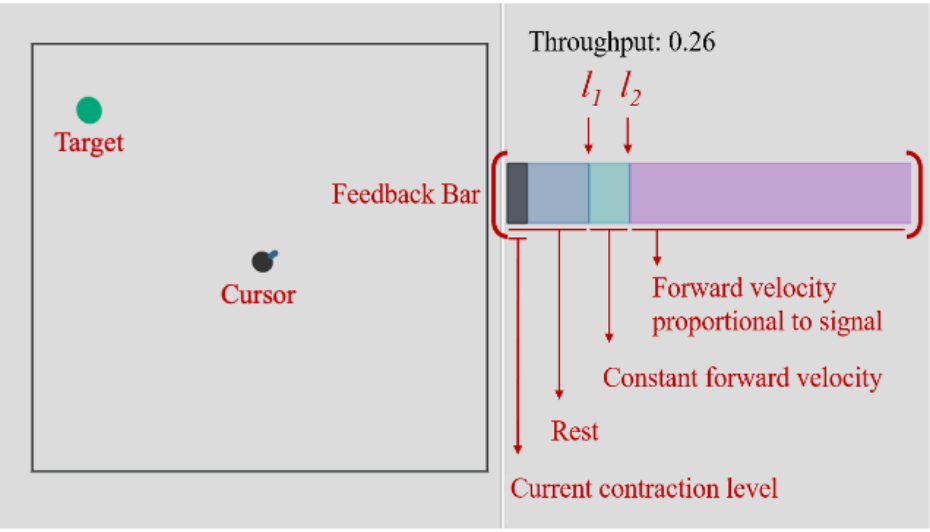



*Figure 2.* User interface for both control methods. The task interface with the cursor and target is shown on the left-hand side. The feedback interface is shown on the right-hand side. The subject controls only the forward motion of the cursor in the auto-rotate method. Both rotation and forward motion are controlled by the subject during the manual rotate method.

### *2.2.3 Cursor Control Method Designs*

The cursor control methods enabled the cursor to move in 2-DOFs, while differing in the level of control afforded to the subject. In the auto-rotate method, the cursor automatically rotated at a set angular velocity and moved forward in the direction of the cursor pointer when the subject contracted the temporalis muscle. In contrast, the subject had direct control over both the rotational and forward motion of the cursor in the manual rotate method. The subject used a muscle contraction pattern to select and control the mode (i.e., rotate or forward).

In the auto-rotate method, the cursor continuously rotated counterclockwise at 120 degrees per second and did not move forward while the input signal was below the first threshold ($\bar{x} \leq l_1 = 0.2$). When the subject's input signal exceeded the first threshold and remained at or below the second threshold ($\bar{x} \leq l_2 = 0.3$), the cursor ceased rotating and moved forward at a constant, minimum velocity, $v_c$, of 0.05 units/s. An input signal above the second threshold, $l_2$, caused the cursor to move forward at a velocity calculated by:

$$v = v_c + v_m(\bar{x} - l_2/1 - l_2)^2. \qquad \textit{Equation 1}$$

In the above equation to calculate the forward cursor velocity ($v$), the variables are defined as follows: $v_c$ is the constant, minimum velocity, $v_m$ is the maximum velocity, $\bar{x}$ is the processed sEMG signal, and $l_2$ is the threshold that separates the minimum, constant velocity from the increasing, proportional velocity. The maximum velocity was 0.55 units/s.

The manual rotate method design employed the same thresholds and velocities as the auto-rotate method design. The subject used a muscle activation pattern to convey intent, and a signal input was any activation above the first threshold. The first signal input selected the mode (rotate or forward) based on the length of the muscle activation, where an activation length of 0.5 s or less selected the rotate mode and a longer activation selected the forward mode. A second signal input within one second of the first signal input then allowed continuous control within the selected mode. The cursor acted according to the selected

mode for the duration of the second signal input, and the cursor mode was indicated with color feedback. Figure 3 depicts the muscle activation pattern for both modes.

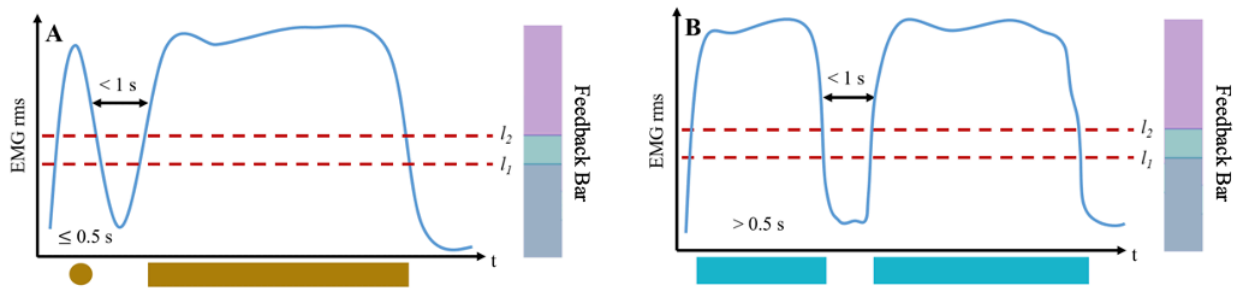



*Figure 3.* Diagram of muscle activation patterns for A) rotate mode and B) forward mode. Each plot indicates the activation times for mode selection (first signal input) followed by mode duration (second signal input, not time limited). The feedback bar from Figure 2 is shown for reference. The Morse code-like pattern is shown under each chart (circle = short signal input, dash = long signal input) and reflects the applicable cursor color feedback. The color feedback stays on for the mode's duration.

### *2.2.4 Analysis*

To measure performance during the Fitts's law-based task, we calculated throughput (TP), overshoots (OS), and path efficiency (PE). TP (bits/s) is a common Fitts's Law task metric for input devices that accounts for movement time (i.e., the time it takes the cursor to reach the target, $MT$) and index of difficulty ($ID$), where $TP = ID/MT$. The OS metric measured the average number of times that the cursor passed through the target. PE was calculated by dividing the straight-line distance from the origin to target by the path length traversed by the cursor. The NASA-TLX workload scores were calculated as described in [56] on a scale of 0 to 100.

## 2.3 Results

### *2.3.1 Throughput, Overshoots, and Path Efficiency*

The results for the performance metrics—TP, OS, and PE—are shown in Figure 4. Generally, subject performance improved across subsequent sessions within a cursor control method. Group 2 appeared to have a higher TP during the training for the auto-rotate method compared to Group 1. Group 2 also

demonstrated higher OS (lower performance) during training for both control methods. A repeated-measures two-tailed t-test on testing data (Session 2 and 3) with Groups 1 and 2 combined showed a significant difference in TP between the auto-rotate and manual rotate methods ($p \ll 0.001$), but no significant differences for PE and OS ($p = 0.128$ and $p = 0.290$, respectively).

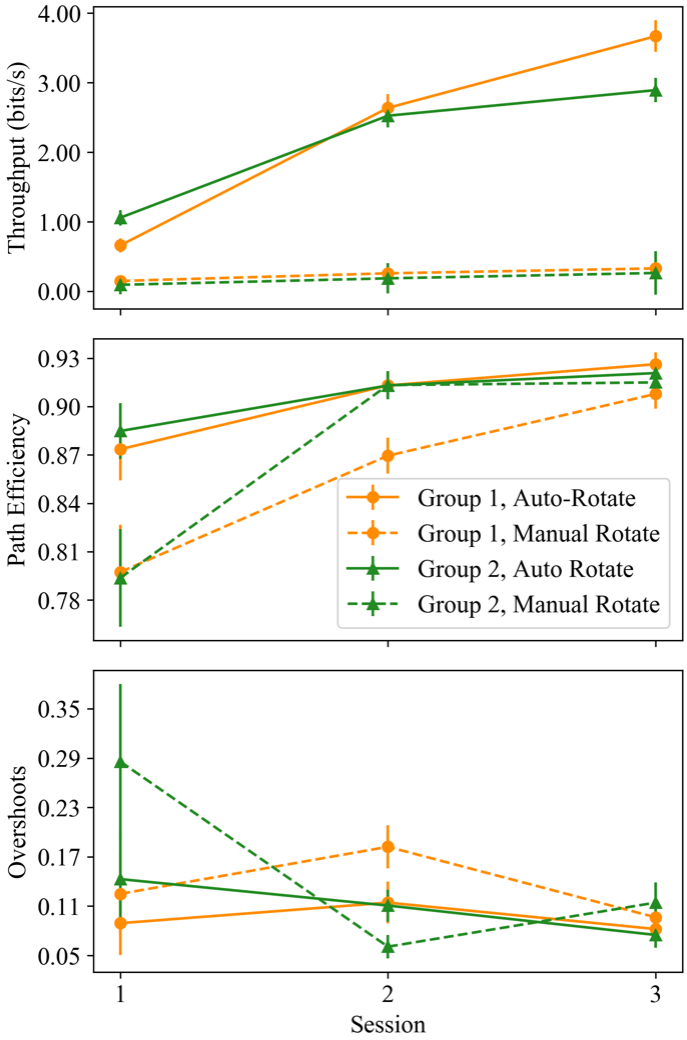



*Figure 4.* Results for Throughput (TP), Path Efficiency (PE), and Overshoots (OS). Session 1 is the training session and Session 2 and 3 are the testing sessions. Error bars denote standard error of the mean.

### *2.3.1 Workload and Cursor Control Method Preferences*

Both Group 1 and Group 2 reported a higher average cognitive workload score for the manual rotate control method, 51.33 ± 13.38 and 40.24 ± 9.15, respectively. In general, Group 2 reported lower cognitive

workload scores with smaller variance for both control methods compared to Group 1. However, the average increase in cognitive workload from the auto-rotate to manual rotate cursor control method was similar (Group 1: 17.19 ± 7.18, Group 2: 14.52 ± 6.80).

One subject's survey was removed from the post-survey analysis because the subject appeared to confuse the cursor control methods. The remaining thirteen subjects all preferred to use the auto-rotate method. Interestingly, the subjects were split on their responses for which cursor control method afforded more control over cursor behavior (7 auto-rotate method, 6 manual rotate method). Two subjects from Group 2 (manual rotate method first) thought that the manual rotate method was more intuitive, and two subjects from Group 1 perceived the manual rotate method as allowing for more precise alignment of the cursor.

### 2.4 Discussion

The results presented for fourteen able-bodied subjects with no prior sEMG experience include performance metrics (TP, PE, OS) and a cognitive workload score for two single-site, noninvasive cursor control methods: auto-rotate and manual rotate. Subjects completed a Fitts's law-based cursor-to-target task that included a training session and two testing sessions for each cursor control method. Subjects were randomly assigned to groups, which determined the order that they evaluated both cursor control methods. Group 1 evaluated the auto-rotate method followed by the manual rotate method, and Group 2 evaluated the methods in the opposite order. Subjects completed pre- and post-session surveys. It was assumed that subjects had adequate control over their temporalis muscle, learned the control method at the conclusion of training, and could comprehend the information displayed on the screen.

The order of the cursor control methods may have affected the overshoots during training for Group 2 (manual rotate first). Group 2 had a relatively high OS average during training for both cursor control methods. It was observed that subjects during the manual control method tended to maintain a basal muscle activation, which may be a result of their need to make more frequent activations to achieve the muscle activation pattern. This may explain the increase in the OS metric during the auto-rotate method training

for Group 2 compared to Group 1, as Group 2 subjects tended not to relax, and in the auto-rotate method the cursor will immediately move forward once the first threshold is exceeded. The higher OS average during training may indicate that these subjects had less control of their sEMG amplitude and less ability to predict the signal decay in order to stop the cursor. Group 2 had a higher TP during the auto-rotate method training, which may be attributed to their accumulated EMG experience aiding their performance for a less complicated control method. This does not seem to be the case for Group 1 when trained on their second control method, manual rotate. However, these differences appear limited to training. The statistical results indicated no differences in the OS and PE metrics during testing, however, TP was found to be significantly higher for the auto-rotate method.

The TP achieved during the auto-rotate method is an order of magnitude higher than the manual rotate method. The performance difference is likely attributed to the MT, which is the time it takes the cursor to reach the target not including reaction or dwell time [55]. For the auto-rotate method, the time starts when the cursor moves forward for the first time. The initial rotation time is not included, which gives the subject the opportunity to line up the target and may increase TP. In contrast, the time starts in the manual rotate method at the first signal input, which is defined as the first time the processed sEMG signal, $\bar{x}$, exceeds the first threshold. During testing the average initial rotation time was 1.29 ± 0.86 s and 1.39 ± 0.85 s, and the average time to first signal input was 0.57 ± 0.40 s and 0.71 ± 0.57 s for Groups 1 and 2, respectively. The average TP was 2.24 bits/s for the auto-rotate method and 0.23 bits/s for the manual rotate method. In comparison, other studies have reported a multi-site EMG TP of 0.84 bits/s [51] and 0.57 bits/s (for a patient) [57]. Studies using multiple muscle sites including the forelimbs reported a TP of 1.3 bits/s [58] and 0.4 bits/s (0.2 bits/s for a patient) [52]. Although our TP results for the manual rotate method were below 1.0 bits/s, the results were on par with similar systems in the literature indicating that these types of myoelectric control systems generally have low communication rates and could be improved.

The PE is high for both groups and may have reflected a collective strategy to precisely hit the targets by accurately lining up the cursor. An alternative strategy would have been to place the cursor near

the target and make alignment corrections when in closer proximity. The apparent choice of this strategy may have been due to the training presentations, where the cursor rotation was described first. The subjects were also told to maximize TP (smaller $MT$ would require higher PE), which may be a factor. It would be interesting to observe if the collective strategy holds with more subjects.

The cognitive workload scores for both groups reflected a similar trend that the manual rotate method induced a higher workload than the auto-rotate one. Both groups had similar increases in cognitive workload scores from auto-rotate to manual rotate, but Group 2 reported lower absolute workload scores for both cursor control methods. Both groups reported similar cognitive workload scores for their first cursor control methods (Group 1: 40.43 ± 6.24, Group 2: 40.24 ± 9.15), which indicated that their initial perception of the workload to use a sEMG-based control system was similar and the second cursor control method was assessed relative to the first even though subjects were instructed to only consider the most recent method.

We hypothesized that the auto-rotate method would result in a higher TP and lower workload, which the results appear to support. The manual control method did not have significantly more precision, as measured by PE. While subjects tended to prefer the auto-rotate method for this target-to-cursor task, approximately half the subjects thought the manual rotate method afforded them more control. Overall, our results suggest that a viable interface for 2D cursor control can be achieved using only a single muscle site, as compared to other studies in the literature that use more than one site.

## 2.5 Conclusions

The results from this study demonstrated the feasibility of a single-site sEMG control method that may be less intuitive than multi-site systems. Lessons learned from this study were used in the design of the subsequent study presented in the next chapter, *Chapter 3*. During the initial testing for this study, user feedback indicated that color feedback on the cursor body aided in understanding the command system. However, the timing and placement of the color feedback was not thoroughly investigated. Furthermore, this initial study only implemented two commands, which required minimal learning. For this sEMG

command system to be extended to more complicated systems, such as a robot that requires more commands, it is necessary to understand how to train the users. Therefore, the study in *Chapter 3* sought to evaluate two different options for color feedback and four training methodologies with the number of commands increased to four.

The study in the next chapter, *Chapter 3,* leveraged the data from the initial study to set system and task parameters. Specifically, the timeout between the first and second inputs was reduced from 1 s to 0.5 s and the maximum trial time decreased from 2 minutes to 1 minute. The durations for short ($\leq 0.5\ s$) and long ($> 0.5\ s$) inputs remained, as they continued to appear reasonable. The feedback bar was removed from the user interface to isolate the forms of sEMG feedback. We expected the 80 trials used in the initial study would be too few for the subsequent study with four commands and doubled the trials to 160. Overall, this initial study provided guidance for developing the next study with increased command complexity.

# Chapter 3

# Myoelectric Control Training Methodologies and Effects on Performance, Workload, and Trust

## Preface

After demonstrating the feasibility of the single-site sEMG myoelectric control methodology established in *Chapter 2*, I wanted to increase the number of commands from two to four to prepare for more complex task control, test different training methodologies, and better understand human interaction with the myoelectric control system. This chapter describes a study that assessed four training methodologies and their effects on performance, cognitive workload, and trust, as well as post-training performance retention. An additional objective was to select the appropriate training methodology for subsequent studies. The contents of this chapter have been previously published in [3] and [2].

## 3.1 Introduction

Increased interest in BoMIs necessitates understanding how to train users to use nontraditional control inputs. The BoMI systems previously studied focused on the development and demonstration of the interface but did not address user training beyond task repetition or test setup familiarization. For a novel input, it is important to efficiently train new users. We decided to apply automation to the skill-based training of learning sEMG control and compare different automated training methodologies. Selection of the training methodology should take into consideration the effects on task and human-automation interaction factors, such as trust and workload, and should also account for context-specific interface elements (e.g., flight displays, communication interfaces).

### *3.1.1 Selected Training Methodologies*

Training novice users to effectively use sEMG control methods can be a long and difficult task [45]. While expert instructors can be effective at increasing learning rates, their time can be costly, and their availability is often limited. For these reasons, we were interested in automated training methodologies which seek to fill the role of an expert instructor. Self-directed practice and instructor-guided practice have both been shown to work for learning complex medical skills [59], though it is unclear how well this paradigm would transfer for non-medical tasks. The use of augmented feedback strategies, however, have been shown to help reduce training times and have the potential to reduce cognitive workload, or the amount of mental effort allocated, in especially demanding tasks [60]. Augmented feedback provides information that "cannot be elaborated without an external source; thus, it is provided by a trainer or a display" and has been shown to effectively improve performance in a wide variety of motor tasks [61]. Augmented feedback has been used for training, motor skill acquisition, rehabilitation, and operational assistance for tasks which range from simple, closed-environment lab demonstrations to complex, operational cases [61]. Recent approaches to using augmented feedback have focused on multimodal cueing and virtual and augmented reality displays in a variety of medical [62]–[64], aerospace, and robotic tasks [65]–[67]. Across these studies, results indicate that providing task-appropriate augmented feedback can improve rate of motor skill acquisition, final performance, and retention, though these results vary depending on the means of the augmented feedback presented. Many of these studies have compared how to provide feedback across different modalities to identify optimal pairings between modality and tasks but investigating when to provide feedback is equally important. Concurrent feedback is presented in real-time, as subjects execute a task, while terminal feedback is presented after the task is completed. Some researchers have found that intermittent, concurrent high-frequency feedback is superior to low-frequency feedback [62], [68], while others have found that terminal feedback outperforms a concurrent approach [69], [70]. It is difficult to generalize between different feedback and task types [71], [72]. In general, however, concurrent feedback has been shown to be more useful with higher functional task complexity, whereas terminal feedback is often less useful when task complexity is high [61]. Concurrent feedback has recently shown great promise

in myoelectric control, though less progress has been made comparing concurrent and terminal feedback strategies or investigating long-term learning effects [73]–[76]. Therefore, it is important to compare concurrent and terminal feedback approaches within the same task to better understand the appropriate application of training methodology.

Biofeedback, which applies augmented feedback strategies to measured physiological signals such as EMG, has proven to be a useful tool for improving performance and assisting in rehabilitation [77]. Researchers have investigated various augmented biofeedback techniques and found that they help subjects to "become more cognizant of their own EMG signal" [78], allowing for better control of their muscle activity. A recent review of the biofeedback literature suggests that "[b]iofeedback is more effective than usual therapy," though they also note that "[f]urther research is required to determine the long-term effect [biofeedback has] on learning" [79]. In our study, we refer to any additional, visual information given about a sEMG signal input or its processed output as "augmented feedback." The augmented training strategies are analogous to alerting automation and the strategy of complementation in aviation [80], where the augmented training strategies cooperate with and provide additional information to the user.

Another possible training methodology that leverages adaptation comes from motor learning. Users learn and adapt to improve performance throughout practice and training, and induced variability can improve performance. For example, Seow et al. varied a task parameter, where subjects either trained on a set thrust level or a variable one for a spaceship video game [81]. Subjects with variable thrust training performed better than those with a consistent thrust when tasked with a novel thrust level [81]. Braun et al. showed during a planar reaching task that randomly varying the feedback uncertainty (i.e., the difference between the actual and shown position of the hand) led to improved skill generalization [82]. Surface EMG control adaptation may follow Bayesian theory [83], where mapping uncertainty—or the uncertainty of the brain's model of the applicable system—increases adaptation rate. Conversely, increased sensorimotor feedback uncertainty decreases adaptation rate. Lyons and Joshi demonstrated that subjects exposed to a mapping uncertainty during cursor control had higher adaptation rates to sensorimotor feedback when the mapping uncertainty was removed, indicating that artificially added mapping uncertainty during training

may increase adaptation rate [84]. Additionally, a noisier sEMG control method (i.e., increased mapping uncertainty) provides more information about the control signal, which can lead to a larger adaptation rate compared to a more filtered classification method [85]. Similarly, introducing noise in a joystick-controlled final approach flight task has been shown to increase the rate of motor skill acquisition [65]. Taken together, these studies indicate that varying a task parameter [81], [82] and increasing mapping uncertainty [84], [85] can potentially improve skill acquisition. However, research in this area is not well-studied for sEMG control or for tasks that do not use the planar reaching paradigm [86]. Therefore, we have included an adaptation-based strategy in our study to compare with more established training methodologies to investigate the feasibility of this method.

### *3.1.2 Workload and Trust in Automation*

When evaluating the different automated training methodologies, it was important to select appropriate training feedback that does not cause additional cognitive demands. Karasinski et al. investigated the effects of concurrent feedback in a simulated, four degree of freedom manually controlled spacecraft inspection task [60]. Subjects in the feedback group performed the task substantially faster and more accurately than those in a control group and reported a significantly lower cognitive workload. Cognitive workload can be defined as "the cost incurred by a human operator to achieve a particular level of performance" [87], [88], and is commonly assessed using quantitative, subjective techniques such as the NASA-TLX [56], [87], [88] and Modified Bedford Workload [89] scales. While there are limits to these subjective techniques, such as the possibility of large intersubject variability, they provide a rapid and easily administered tool for cognitive workload assessment. The Modified Bedford Workload scale was administered frequently throughout this study as a rapid means of estimating spare cognitive capacity. In addition to subjective cognitive workload measures, there are a variety of techniques which aim to objectively estimate cognitive workload. Physiological measurements such as heart rate variability have been used to assess workload for many types of actual and simulated aerospace flight tasks [90]. Heart rate and heart rate variability also provide objective measurements that have been previously correlated with subjective scales such as Bedford

and NASA-TLX [91]. Researchers have had mixed success in being able to generally relate physiological measures to subjective workload scales, however, and physiological workload estimates were not used in this study. Among the most common objective measurement techniques, however, is the secondary task, which requires subjects to complete the primary task, and then use any spare cognitive margin to respond to an additional task [92]. This secondary task approach does not work well for simpler tasks, however, as completing it begins to compete for attention with the primary task [93]. No objective workload measurements were included in this study.

Trust is another important factor when considering human-automation interaction, and inappropriate trust can lead to the disuse or misuse of automated systems. Trust in an agent is "the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability" [94]. The reliability of a system, in particular, has been shown to be an important aspect of an operator's trust in a system [95]. While there have been many proposed models for trust, Hoff and Bashir's three layer model deserves particular attention [96]. After performing a systematic review of the literature, they developed a three-layer model of trust which is split between dispositional trust, situational trust, and learned trust. Dispositional trust is an individual's tendency to trust automation, which is a relatively stable trait compared to situational and learned trust [96]. Situational trust is influenced by external and internal factors, such as the system design and cognitive workload [96]. Learned trust is impacted by the user's experience and perception from interacting with the specific system [96]. Though researchers have little control over dispositional trust, they can affect situational trust by varying an experimental interface or environment and learned trust can be evaluated using repeated measures. Our study aimed to alter situational trust through different automated training methodologies, where the agent (i.e., the automated feedback in place of an instructor) helped the subjects learn a skill. We expected subjects in each group to perceive whether the agent will help them achieve their goals differently based on the assigned training methodology. We also assessed learned trust with measurements taken periodically during the study. The subjects did not have an option to turn off or alter the agent but may have felt that the agent

was not aiding in their goals. If subjects became reliant on the agent, we expected to see decreased performance when the agent was removed.

### *3.1.3 Study Objective*

While it is not the objective of this work to develop or optimize a sEMG system for a particular application, the sEMG system can act as a testbed in which the effects of different automated training methodologies are evaluated, and the novelty in using these systems enables the observation of early learning effects. The sEMG-controlled Fitts's law-based task from *Chapter 2* naturally serves as an excellent testbed for evaluating augmented feedback strategies. Most participants cannot easily predict or quickly understand the signal they generate and pass into the sEMG controller, so augmented feedback can provide insight into what would otherwise be an inherently noisy and elusive process. Current systems increasingly incorporate humans with automation, necessitating an enhanced understanding of human-automation interaction. The purpose of this study was to investigate human-automation interaction with an emphasis on performance, workload, and trust during early learning across different automated training methodologies.

To our knowledge, there has been no previous study investigating differences in performance, workload, and trust across different automated training methodologies using augmented feedback and adaptation-based method. In this study, we addressed the effects of automated training methodology on performance, cognitive workload, and trust during a computer-based Fitts's law-based [97] cursor-to-target task. The training methodologies include repetition, concurrent feedback, terminal feedback, and an adaptation-based method. The treatments are removed for the evaluation phase to assess if subjects in the augmented feedback conditions—concurrent feedback and terminal feedback—succumbed to the guidance hypothesis [98], which would indicate that they have become reliant on the feedback in order to perform the task. The experiment was designed to provide subjects with a sufficiently challenging task to observe early learning effects and the evolution of performance, cognitive workload, and trust.

## 3.2 Materials and Methods

### *3.2.1 Subjects and Experiment Setup*

The UC Davis Institutional Review Board approved the study protocol, and subjects were recruited from the university student population. Exclusion criteria for subjects included a history of neuromuscular disorders, physical limitations of dominant arm, and prior myoelectric control experience. Subjects provided written consent prior to participation. A total of 55 subjects volunteered, and subjects were released from the study due to equipment issues ($N = 3$), withdrawal request ($N = 2$), and significant motivation issues (e.g., not attempting the task; $N = 2$). The remaining 48 subjects completed the protocol and had an average age of 20.1 ± 1.4 years ($\mu \pm \sigma$), included 2 left-hand dominant subjects, and had an equal participation of men and women.

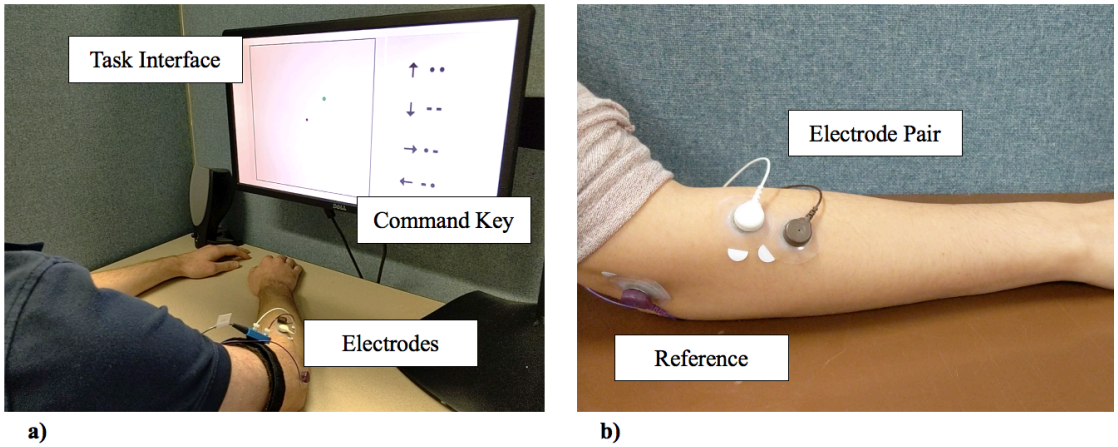



*Figure 5.* Setup with interface displayed during training and testing of cursor-to-target task (a) and electrode placement (b).

Subjects were trained to control a cursor using sEMG to perform a Fitts's law-based cursor-to-target task with a center-out paradigm [1], [52]. During the experiment, subjects sat in front of a desk and a computer screen (see Figure 5a) with electrodes on their forearms. Two electrodes (ConMed 1620 Ag/AgCl center snap) approximately 2.5 cm apart were placed on the dominant hand side near the extensor digitorum proximal attachment. A reference electrode was located near the lateral epicondyle of the

humerus (see Figure 5b). The electrodes' signal was acquired as described in [54], the signal processing followed [1], and the experimental software used the UC Davis-developed Python AxoPy library [99]. The rms value for each time-domain sample window was calculated, normalized by a manually set calibration constant (in mV), and incorporated into a 0.5 s moving average window to yield an updated and processed sEMG signal, $\bar{x}$ (mV/mV), at 16 Hz (used in Equation 1). The subjects learned to manipulate their sEMG signals to produce serial patterns that translated to cursor motion.

After the electrode placement, the subjects viewed their sEMG signal on an oscilloscope shown on the computer screen and were instructed to flex their hand to induce signal changes. This oscilloscope activity confirmed proper electrode adhesion and illustrated the subjects' abilities to intentionally change their sEMG signal. For the remainder of the study, the subjects individually determined how to produce sufficient muscle activation and were prompted to try different, self-selected movements and contractions during the manual calibration. Subjects self-reported the muscle activation strategies they used at the end of the study. Flexing the wrist was the most common strategy (*N = 34*), possibly due to its introduction during the oscilloscope activity. Other strategies included making a fist, raising a single finger, or raising multiple fingers. Most subjects (*N = 33*) only used one strategy. Fifteen subjects elected to use multiple strategies, and five of those subjects used specific strategies for different actions (e.g., raising fingers for shorter inputs). A summary of the strategies used by experimental group is shown in Table 1. A more detailed discussion of the experimental groups is provided in section *3.2.3 Command Design and Group Treatments*.

*Table 1.* Summary of Muscle Activation Strategies by Group

| Strategy, N | Group | | | | |
|---|---|---|---|---|---|
| | Control | Concurrent Feedback | Terminal Feedback | Adaptive Threshold | All |
| Flex Wrist | 7 | 10 | 9 | 8 | 34 |
| Make Fist | 5 | 3 | 2 | 3 | 13 |
| Raise Finger | 1 | 0 | 2 | 1 | 4 |
| Raise Multiple Fingers | 3 | 6 | 5 | 6 | 20 |
| One Strategy | 9 | 8 | 8 | 8 | 33 |
| Multiple Strategies | 3 | 4 | 4 | 4 | 15 |

### *3.2.2 Cursor-to-Target Task*

In our study, the Fitts's law-based task served to provide an interactive, engaging environment in which the subjects developed their proficiency of sEMG commands by moving a cursor to hit a target. The Fitts's law-based task used a center-out paradigm and a similar interface as in [1]. The square cursor interface had normalized horizontal and vertical bounds of [-1,1] and a length of 2 units. Each trial began with the cursor at the center of the interface and a target in a pseudorandom position. There were 40 unique target positions that covered a range of index of difficulties (*IDs*) from 1.00 to 4.09 bits (calculated by the Shannon formulation [55]). The subject used sEMG to convey commands and moved the cursor to the target; the cursor had to dwell on the target for 1 s to successfully complete the trial. The maximum trial time was 60 s, which was determined by reviewing previous data [1] and preliminary testing.

For myoelectric control, it is necessary to designate a threshold below which the signal input is considered at rest because sEMG signals effectively always have non-zero values. The threshold, $l_1$ (in mV/mV), defined the crossover value for a signal at "rest" versus active (see Figure7a). From our previous observations, cursor control improved when the initial motion had a slow, constant velocity before switching to a velocity proportional to the signal input, as it allowed subjects to maintain a slow velocity when desired (e.g., in close vicinity to a target). Therefore, we designated a second threshold, $l_2$ (in mV/mV), that delineated the constant from the proportional velocity control. The cursor was either stationary ($\bar{x} < l_1$), moving at a small, constant velocity ($l_1 < \bar{x} \leq l_2$), or moving at a velocity proportional to the input ($\bar{x} > l_2$). When the input exceeded $l_2$, the cursor velocity was calculated by Equation 1, where $v_c$ is the minimum velocity (0.05 units/second), $v_m$ is the maximum velocity (0.50 units/second), $l_2$ is 0.30 mV/mV, and $\bar{x}$ is the filtered, averaged sEMG signal. The $l_1$ value was nominally 0.20 mV/mV, except for the Adaptive Threshold group in which the value was randomly selected for each trial ($l_1$ = 0.10, 0.15, 0.20, 0.25, 0.30 mV/mV).

Since the subjects could only select one command at a time, the cursor could either move up, down, left, or right. It was not possible to combine commands, which restricted the cursor motion to a rectilinear trajectory. To illustrate the resulting cursor trajectories, Figure 6 displays selected trials that are

either within one standard deviation of the median (see Figure 6a) or outside of one standard deviation (see Figure 6b) for a path efficiency metric. Example trajectories are superimposed on a cursor position heat map of successful trials for the applicable target (*ID* = 4.09). The trajectories and heat maps are from across groups for the selected target position.

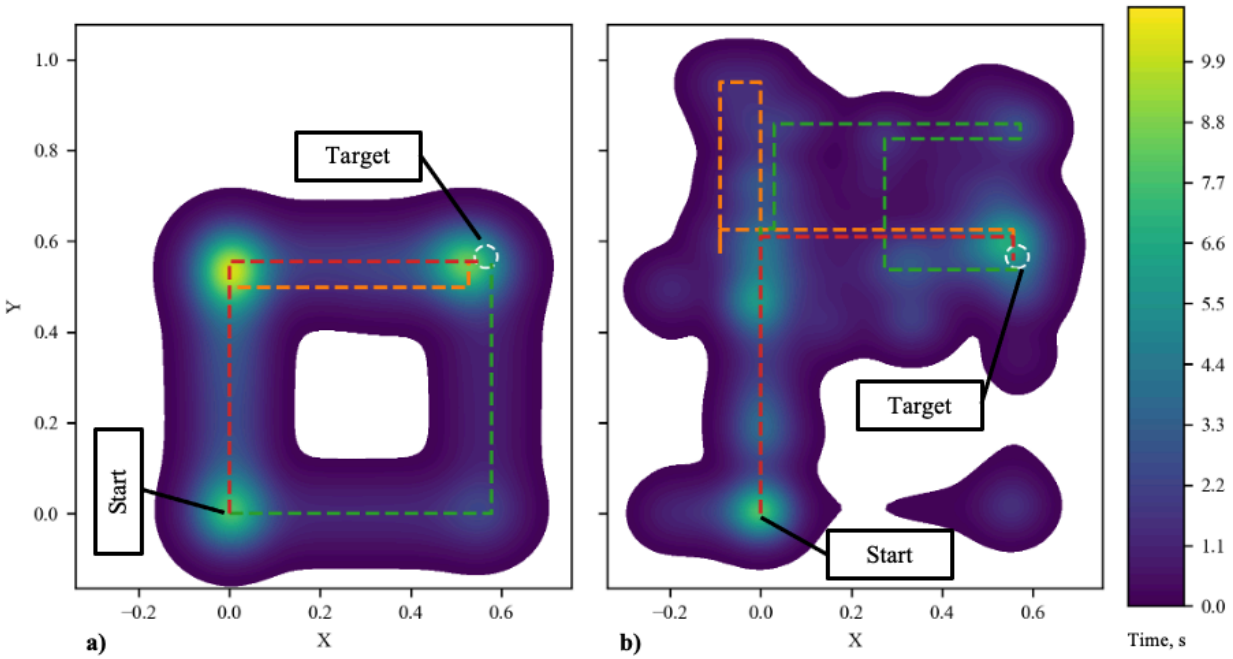



*Figure 6.* Sample cursor trajectories for a target (*ID* = 4.09) either a) within or b) outside one standard deviation of the median for path efficiency. Each dashed line shows a trajectory from a subject. Dashed line colors are varied to differentiate between individual trajectories. The target outline is shown with a white, dashed line.

### *3.2.3 Command Design and Group Treatments*

The premise of the sEMG control methodology was to use serial patterns of muscle activation ("inputs") to convey commands, similar to Morse code [53]. In our scheme, the first two inputs selected the command and a third input allowed for continuous control of forward movement in the selected direction (see Figure7a). A timeout between the first two inputs allowed a reset in the case of errors during command selection. The command inputs were defined by the duration that the sEMG signal exceeded the threshold, $l_1$, and each input was identified as "short" ($\leq$ 0.50 s) or "long" ($>$ 0.50 s). For example, the combination of two "short" inputs selected the "up" command. Subjects learned the 2-input code for four commands:

up, down, left, right. A portion of the user interface contained a command key so that the subjects did not need to memorize the serial, sEMG patterns to produce commands (see Figure 7c).

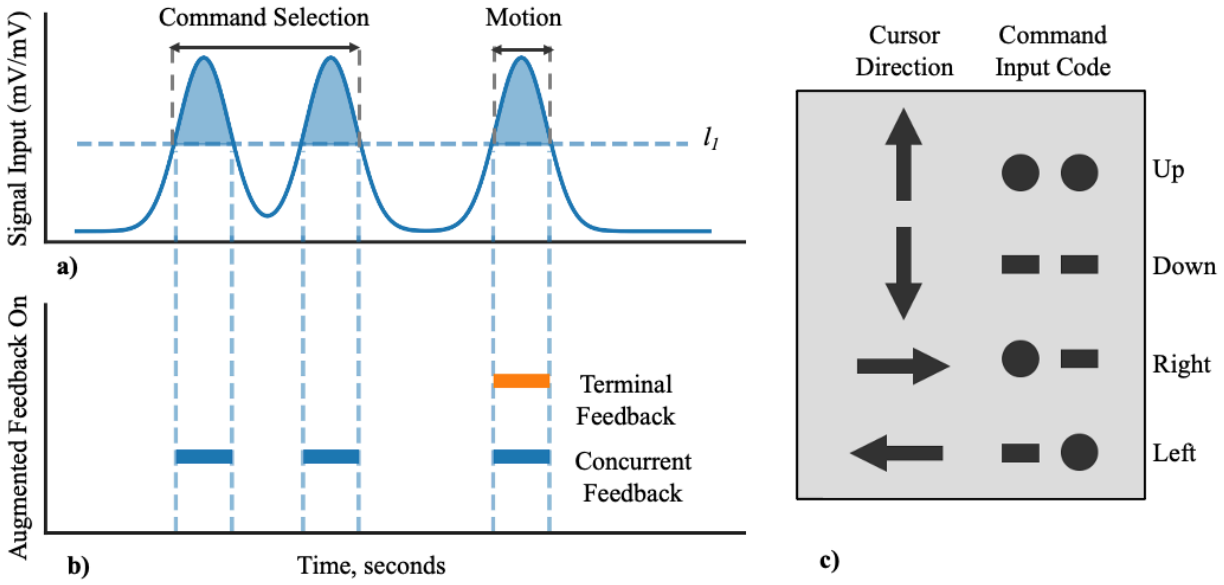



*Figure 7.* Illustrative signal input (a) with the feedback timing (b) for applicable groups. The command key (c) is always present (dot = short, dash = long)

The four groups were each assigned one training methodology:

1) The Control group trained solely through task repetition and were able to view the motion of the cursor.
2) The Concurrent Feedback group received additional visual feedback during the task that indicated when the processed sEMG signal, $\bar{x}$, exceeded the threshold, $l_1 = 0.20$ mV/mV, which was important for selecting commands. The subjects received feedback to confirm reception of the signal by the cursor changing color (see Figure 7b for timing and Figure 8b for coloring).
3) The Terminal Feedback group received visual feedback after entering a command (see Figure 7b). The subjects received feedback regarding the interpretation of the signal to a command and signal reception during motion.
4) The Adaptive Threshold group was the same visually as the Control group, but the threshold, $l_1$, was randomly selected at each trial among values between 0.10 and 0.30. In this group, the processed sEMG

signal, $\bar{x}$, had to exceed the threshold that changed each trial to input commands. These subjects trained with variable parameters/increased mapping uncertainty (i.e., whether their signal will cross the threshold) with the aim of improving motor skill acquisition.

As shown in Figure 8, all the groups viewed a similar interface. Our previous findings in providing augmented feedback for manual control tasks found that visual augmented feedback worked best when placed on an element of the display that operators were already visually engaged with, and that visual feedback placed elsewhere would both lower performance and increase workload [100]. This takes advantage of operators' normal patterns of visual attention, which is "drawn to display items that are…, colorful, and changing (e.g., blinking)" [95]. For this study, we sought to identify differences between providing feedback concurrently (every time subjects crossed $l_1$), and terminally (only after they had successfully input a command). Additionally, we selected the feedback colors to be color-blind agnostic. The additional visual feedback in the Concurrent Feedback group was provided by changing the color of the cursor (Figure 8b). The Terminal Feedback group had additional visual elements at the edges of the cursor interface that changed color to reflect the command (Figure 8c). The visual element locations were selected based on pilot testing. Both of the two augmented feedback techniques provided users with feedback when the system had received a valid input (see the colored regions for the timing shown in Figure 7b), and the Terminal Feedback group received the additional feedback of the system's interpretation of the direction to move the cursor. A recent review of augmented feedback techniques (see Ref. [61]) found that terminal feedback is more effective for tasks with low functional complexity, while concurrent feedback is more effective for tasks with high functional complexity. A summary of the training type, displays, and visual feedback are provided in Table 2 (Refer to *Appendix B* for a storyboard walkthrough of each interface). It should be noted that the subjects did not have a choice to disuse or turn off the automated training methodology. For example, this is similar to the predetermined automation design of a space exploration autonomous behavioral health tool reported in [101].

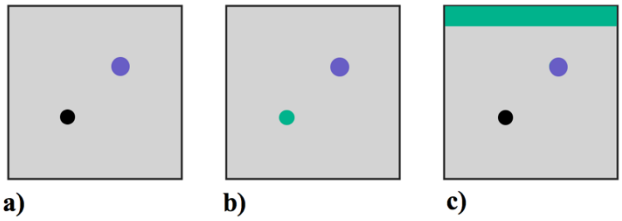



*Figure 8.* Illustrative cursor interfaces for a) the Control and Adaptive Threshold, b) Concurrent Feedback and c) Terminal Feedback groups (not to scale). The target is shown as a purple circle. The cursor appears as a black circle for the Control, Terminal Feedback, and Adaptive Threshold groups. For Concurrent Feedback group, the cursor appears as a black circle when the signal is below the threshold, $l_1$, and green circle when the signal is above the threshold (b). In the Terminal Feedback group, a green bar appears at the edge of the interface to indicate the selected command. The example shown in (c) indicates that the up command has been selected. (Selecting the left command would cause the green bar to appear at the left edge and so on.)

*Table 2.* Summary of Group Treatments and Displays

| Group | Type | Display | Visual Feedback Timing |
|---|---|---|---|
| Control | Repetition | Figure 8a | N/A |
| Concurrent Feedback | Augmented Feedback | Figure 8b | Figure 7b |
| Terminal Feedback | Augmented Feedback | Figure 8c | Figure 7b |
| Adaptive Threshold | Motor Adaptation | Figure 8a | N/A |

Subjects were assigned to one of four groups (12 subjects per group) and proceeded through the same protocol. The protocol consisted of the following general steps: consent, entrance survey, experiment overview and setup, pre-training assessments, training phase, post-training assessments, evaluation phase, post-evaluation assessments, and an exit survey. A maximum voluntary muscle contraction measurement was collected pre-training and post-evaluation to confirm that subjects were not significantly fatigued during the experiment, and which has been shown to affect sEMG signals [26].

We developed a "Command Accuracy Test" to assess each subject's ability to produce commands. The Command Accuracy Test was conducted pre-training, post-training, and post-evaluation to provide an assessment similar to offline classification accuracy (see Ref. [102] for a sEMG example). A Fitts's law-based task was implemented for the training and evaluation phases. There was a total of 160 trials with the first 120 trials used for training and the last 40 trials for evaluation. Trials were grouped into sets of 10 to form a Block; there was a 30 s minimum rest time after each Block during which subjects completed surveys for workload and trust.

## 3.3 Analysis and Hypotheses

The analysis focused primarily on command and interaction (i.e., cognitive workload and trust) metrics to observe the effects of the automated training methodologies. Task-based metrics were also analyzed. A brief description of the analysis is provided in the following subsections.

### *3.3.1 Command Metrics*

The Command Accuracy Test provided an opportunity to evaluate subjects' proficiency in producing desired commands outside of the Fitts's law-based task and was administered pre-training, post-training, and post-evaluation. During the Command Accuracy Test, subjects responded to a prompt to produce a specified command (either up, down, left, or right) and performed each command 5 times for a total of 20 commands in a pseudorandomized order. The command was scored as successful when the input matched the prompt; inputs classified as any other command were not considered successful. The percent of successful commands was calculated for each subject and averaged within the group for a command accuracy metric. For all the successful commands, command time and average command amplitude were also calculated. The command time was the duration between the start and end of a command input (see Figure 7a). The average command amplitude was the average value of the input above the threshold during the command input (i.e., the shaded regions in first two inputs of Figure 7a); an average command amplitude close to zero represents an efficient input. Chhabra and Jacobs [103] used a similar signal magnitude metric to evaluate input proficiency. These metrics assessed the accuracy and efficiency of command input but were not calculated during the Fitts's law-based task because subject intention (i.e., desired command) would need to be inferred for most time points.

### *3.3.2 Interaction Metrics: Cognitive Workload and Trust*

Cognitive workload was measured using Modified Bedford Workload scale [89]; the subject rated their cognitive workload using a flow diagram (see *Appendix C*). Trust was evaluated using Jian et al.'s twelve

statement questionnaire which measures trust between people and automated systems [104]. Each of the twelve statements was evaluated on a seven-point Likert-type scale. Subjects completed these surveys after each Block in the training and evaluation phases. A Block consisted of 10 trials, where a trial is a single Fitts's law-based cursor-to-target task.

### *3.3.3 Cursor-to-Target Task Metrics*

The task metrics from the training and evaluation phases included percent of successful trials, completion time, throughput (TP), and normalized path length (nPL). The percent of successful trials consisted of the number of successful trials out of the 10 trials within a Block, and only successful trials contributed to the completion time, TP, and nPL metrics. Completion time was defined as the time from the start of the trial to the trial completion but does not include the 1 s dwell time. TP is a common metric for evaluating Fitts's Law tasks and is defined as $TP = ID/MT$, where $ID$ is the index of difficulty and $MT$ is the movement time. The ID is calculated by the Shannon formulation [55] and accounts for the distance between the cursor and target, as well as the target diameter. The MT is time from cursor motion onset to target selection (excluding 1 s dwell time). Given the control scheme and the target positions, there was an optimal, rectilinear path that minimized the distanced travelled by the cursor to select the target. The nPL metric assessed the efficiency of the path traveled by using the optimal path distance to normalize each value. There were 10 unique optimal paths due to the radial symmetry of the target positions.

The percent of successful trials was evaluated every Block. The 40 unique target positions repeated every four Blocks (a.k.a. a Session), so the completion time and nPL metrics were averaged over every four Blocks. We found completion time and nPL to be sensitive to target location or $ID$, therefore averaging over four Blocks (i.e., the 40 unique target positions) was more representative of overall performance. TP is intended to evaluate an input device's efficacy under the assumption that the user is proficient. Therefore, TP was only calculated during the evaluation phase when subjects had reached a performance plateau.

### *3.3.4 Hypotheses*

Based on our prior experience with augmented feedback and myoelectric control, we formed the following hypotheses:

1) The Adaptive Threshold group will have the highest command proficiency by the end, followed by the Concurrent Feedback and Terminal Feedback groups, and then the Control. Varying parameters (i.e., thresholds) would improve sensitivity to the sEMG dynamics. Augmented feedback would provide better sensitivity to the sEMG dynamics than the Control.
2) The Concurrent Feedback and Terminal Feedback groups will have a high level of trust during training with some slight decrease during evaluation. The Control group's trust will continually increase. The Adaptive Threshold group will have lower trust during training, which will increase in the evaluation phase.
3) The workload will continually decrease during the training phase for all groups with the largest decreases for the Concurrent Feedback and Terminal Feedback groups. There will be no significant difference in workload in the evaluation phase for all groups since the training phase was sufficiently long for all groups to proficiently learn to execute the Fitts's law-based task.
4) During the training phase, the Concurrent Feedback group will have the highest performance followed by Terminal Feedback, then Control, and finally the Adaptive Threshold groups. All groups will perform similarly in the evaluation phase since the training phase was sufficiently long for all groups to proficiently execute the Fitts's law-based task.

## 3.4 Results

Subjects were evenly divided into four groups: Control, Concurrent Feedback (visual feedback when $\bar{x} > l_1$), Terminal Feedback (visual feedback after command selected), and Adaptive Threshold ($l_1$ varied on a trial-by-trial basis). Subjects completed a total of 16 Blocks, each comprised of 10 cursor-to-target trials, for a total of 160 trials over the course of the study. When applicable, sets of four Blocks were grouped into

a Session. Each Session was identical, such that they contained the same sequence of pseudorandomized target positions. Muscle fatigue did not appear to alter the results, as the maximum voluntary contraction was not significantly different between the beginning and end of the study ($F(1, 46) = 2.554, p = 0.117$). One subject was removed from this analysis as they changed their arm position/contraction method during testing, which affects the signal.

We ran two-factor linear mixed effect models to investigate changes in command and task performance, workload, and trust with one between-subjects factor, Group, and one within-subjects repeated measure, Block/Session/Test. When significant effects were observed, post hoc comparisons using the Tukey Honest Significance Difference (HSD) test were performed and considered significant at the $p < 0.05$ level, and the Satterthwaite method was used to calculate the adjusted degrees of freedom using the lmerTest package in R [105].

### *3.4.1 Command Metrics*

The results in this section—command accuracy, command time, average command amplitude—are assessed from the Command Accuracy Test. The Command Accuracy Test occurred three times (before training, after training, and after evaluation), and the average was calculated within a Group for each Test. We did not find evidence that analyzing results by command type (i.e., up, down, left, or right) altered the overall findings.

Subjects responded to command prompts in each Command Accuracy Test. The command accuracy indicates the percentage of the 20 prompts in each Test that subjects correctly performed (see Figure 9). There was a significant main factor of Test ($F(2, 88) = 108.485, p < 0.001$), but Group was not significant ($F(3, 44) = 2.631, p = 0.062$). The interaction effect between Group and Test was not significant ($F(6, 88) = 0.826, p = 0.553$). Investigation into the Test variable showed that subjects performed significantly better between Test 1 and 2, and between Test 1 and 3, but not between Test 2 and 3. These results demonstrated that there was a significant improvement in the percent of accurate commands by the end of the training phase.

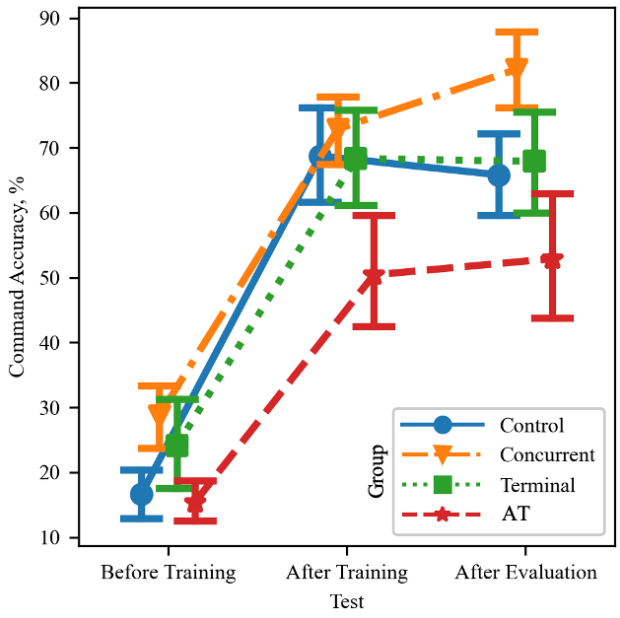



*Figure 9.* Command Accuracy Results. Test 1, 2, and 3 occurred prior to the training phase, after the training phase, and after the evaluation phase, respectively. The Adaptive Threshold Group is abbreviated as "AT."

The command time and average command amplitude metrics were calculated only for successful commands. For command time, there was a significant factor of Test ($F(2, 2055.358) = 9.417, p < 0.001$), but Group ($F(3, 44.080) = 0.662, p = 0.580$), and the interaction between Group and Test ($F(6, 2054.727) = 1.005, p = 0.420$) were not significant. Examining the Test variable, the command times significantly improved from Test 1 to Test 2 and Test 1 to Test 3, but not from Test 2 to Test 3 (see Figure 10). The average command times across Groups were 1.81 s, 1.69 s, and 1.63 s in order of Test.

Average command amplitude had a significant interaction factor of Group and Test ($F(6, 2044.772) = 17.143, p < 0.001$), and the main factor of Test ($F(2, 2044.882) = 13.886, p < 0.001$), but not Group ($F(3, 44.211) = 1.062, p = 0.375$). The only significant Group and Test interactions occurred in Test 1 between the Control and Adaptive Threshold groups and the Concurrent Feedback and Adaptive Threshold groups. Notably, the Adaptive Threshold group's average command amplitude increased from Test 1, whereas all other groups generally declined (see Figure 11). An average command amplitude close to zero represents an efficient input.

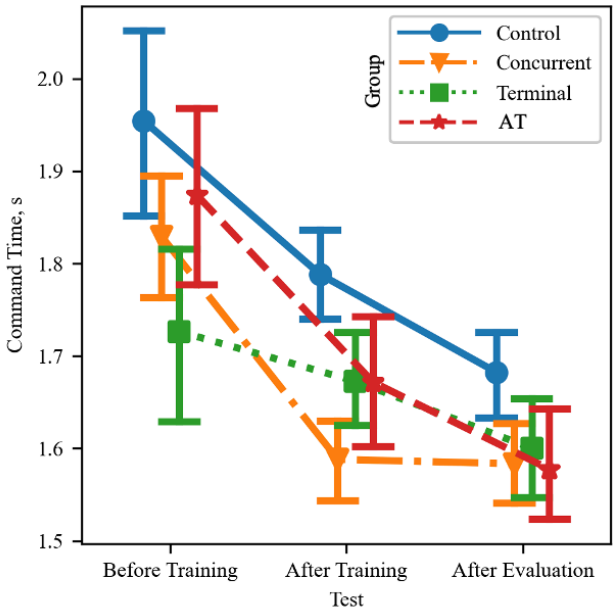



*Figure 10.* Command Time Results. Test 1, 2, and 3 occurred prior to the training phase, after the training phase, and after the evaluation phase, respectively. The Adaptive Threshold Group is abbreviated as "AT."

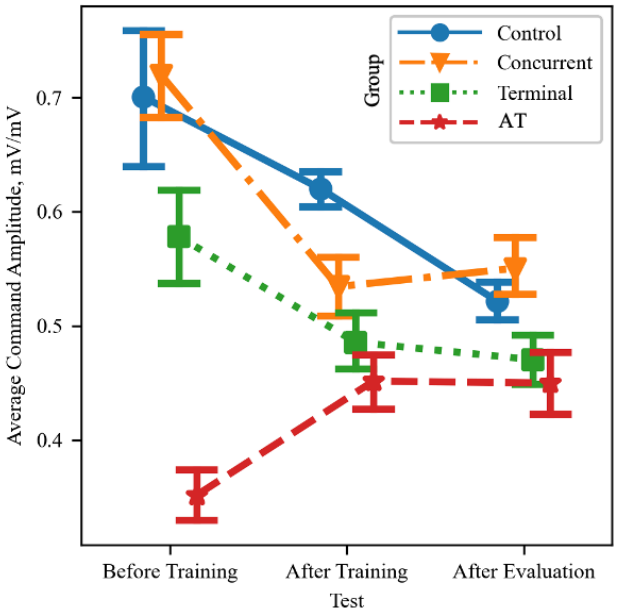



*Figure 11.* Average Command Amplitude. Test 1, 2, and 3 occurred prior to the training phase, after the training phase, and after the evaluation phase, respectively. The Adaptive Threshold Group is abbreviated as "AT."

### *3.4.2 Interaction Metrics: Cognitive Workload and Trust*

The study assessed the changes in cognitive workload—an important consideration for human-automation interaction. The Modified Bedford Workload metric is a subjective measurement of cognitive workload

that ranges from 1-10, where 1 indicates low workload and 10 indicates high workload. There was a significant main factor of Block ($F(15, 660) = 18.284, p < 0.001$), but Group was not found to be significant ($F(3, 44) = 2.164, p = 0.106$). There was also a significant interaction effect between Group and Block ($F(45, 660) = 1.818, p = 0.001$) (see Figure 12). The interaction effect resulted from subjects reporting lower cognitive workload as they learned the task at different rates, indicated by the Block factor. In further investigation of the interaction, we observed that the Adaptive Threshold group reported a significantly higher cognitive workload than the Concurrent Feedback group for Blocks 9, 10, and 11. This perception of high cognitive workload may have resulted from the uncertainties of the changing threshold. None of the groups showed a significant difference in cognitive workload compared to the Control group, and all four groups reported statistically similar workloads in the evaluation phase.

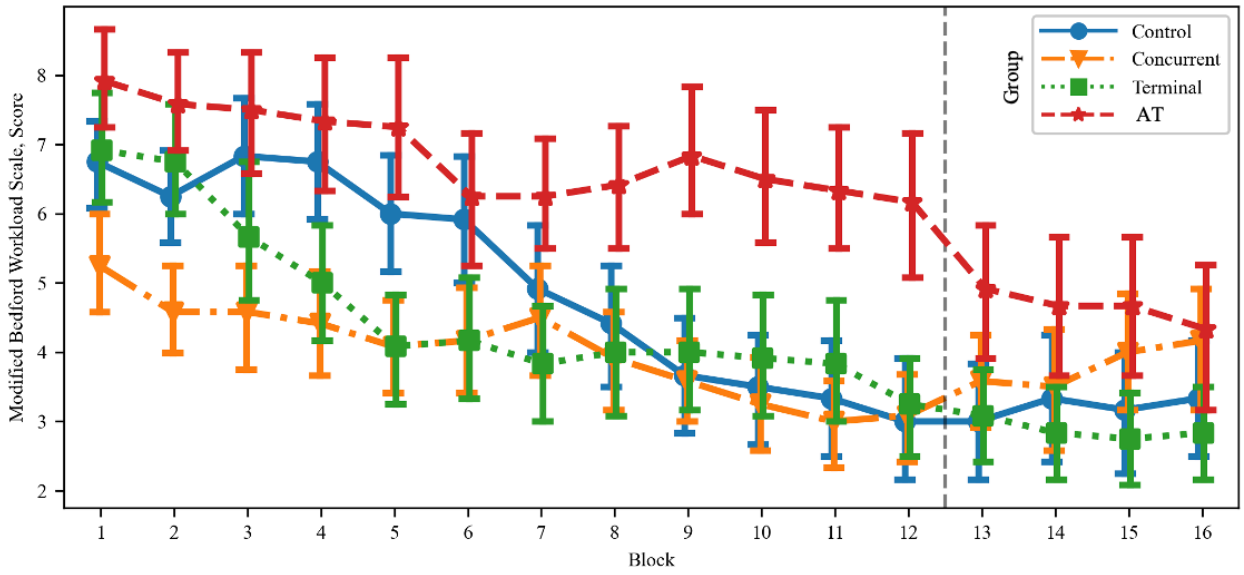



*Figure 12.* Modified Bedford Workload Score. The vertical dashed line represents the transition from training to evaluation. Error bars shown are the standard error of the mean. The Adaptive Threshold Group is abbreviated as "AT."

Intragroup changes in workload are also of interest. The Concurrent Feedback group showed no statistically significant changes in workload between Blocks, though they did demonstrate a nonsignificant, increasing workload trend in the evaluation phase of the experiment. The Terminal Feedback group had statistically higher initial workload for Blocks 1, 2, and 3, but the remainder of the Blocks had a statistically

similar level of workload to each other. The Control group's workload was significantly higher for Blocks 1-6 but leveled off for the remainder of the trials. Finally, the Adaptive Threshold group reported the highest workload in Blocks 1-5, but also saw the largest improvement transitioning into the evaluation phase where their threshold, $l_1$, stabilized to the same fixed value as the other groups.

The second human-automation interaction factor evaluated in the study was trust. Trust was measured using Jian et al.'s twelve question trust in automation survey [104], and administered at the end of each Block. Each question was rated on a 7-point Likert scale, the five reverse coded questions were reversed[1], and the results were averaged to create a single trust score (see Figure 13). There was a significant main factor of Block ($F(15, 660) = 13.051, p < 0.001$), but Group was not found to be significant ($F(3, 44) = 2.588, p = 0.065$). There was also significant interaction effect between Group and Block ($F(45, 660) = 1.996, p < 0.001$). The significant main effect of Block shows a gradual increase in trust throughout the duration of the study. After investigating the interaction effect, we saw that no group reported a significantly different trust level than the Control group on any Block, but that the Adaptive Threshold group recorded a significantly lower trust than the Concurrent Feedback group on Blocks 3-6 and 9. Similar to cognitive workload, the primary interaction effects appeared driven by intragroup differences. The Concurrent Feedback group showed no statistically significant changes throughout the study, the Terminal Feedback and Control groups displayed significant increases in trust in Blocks 1-6, and the Adaptive Threshold group reported significantly higher trust in the evaluation Session than during early Blocks. We investigated the ratings for each individual question in the trust survey but did not observe any trends that contradicted the overall trust score patterns shown in Figure 13.

[1] These questions asked about distrust in the system and the subject rated the system from 1 (Not at all) to 7 (Extremely), which was the same scale for the question asking about trust. Reverse coding means a rating of 7 for a distrust question would become a 1 in order to calculate an overall trust score.

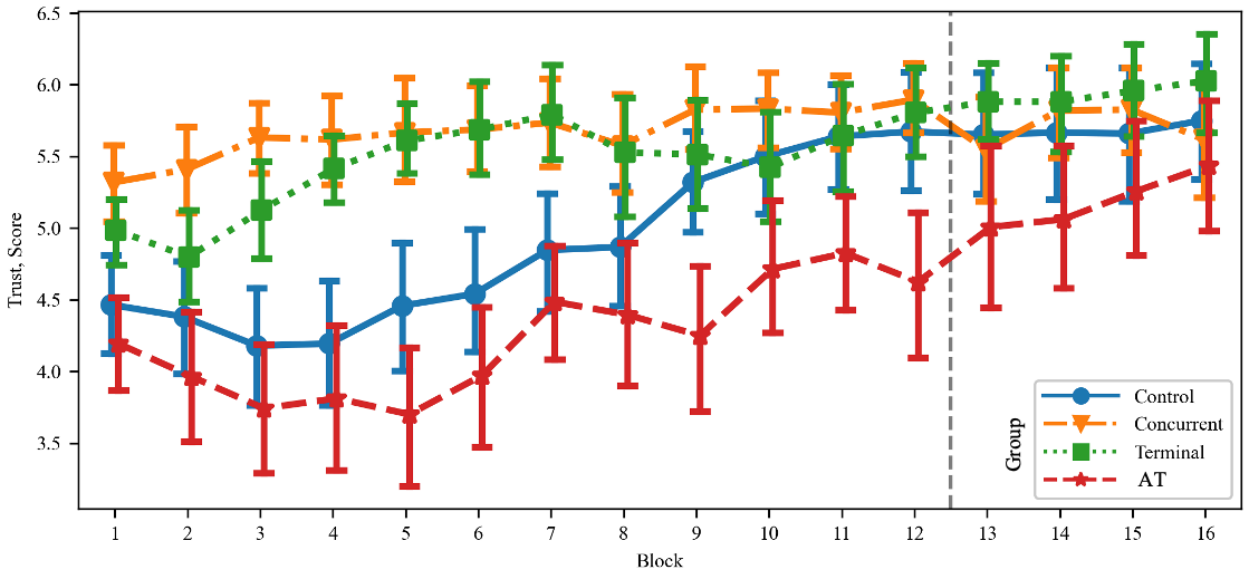



*Figure 13.* Trust Score. The vertical dashed line represents the transition from training to evaluation. Error bars shown are the standard error of the mean. The Adaptive Threshold Group is abbreviated as "AT."

### *3.4.2 Task Metrics*

Although the command metrics were the primary proficiency indicator, we did evaluate aspects of task performance including percent of successful trials, TP, completion time, and nPL. The percent of successful trials metric measured the percentage of successfully completed trials within a Block. For the percent of successful trials, there were significant main factors of Group ($F(3, 44) = 8.183, p < 0.001$) and Block ($F(15, 660) = 31.805, p < 0.001$). There was also a significant interaction effect between Group and Block ($F(45, 660) = 3.903, p < 0.001$). Despite the presence of an interaction effect that resulted from subjects learning the task (as indicated by the Block factor), the main effect of Group could still be interpreted. A Tukey test showed that the subjects in the groups differed significantly, with subjects in the Concurrent Feedback group performing significantly better than those in the Control group ($p = 0.020$). The Tukey test also showed that subjects in the Adaptive Threshold group performed significantly worse than those in the Terminal Feedback and Concurrent Feedback groups ($p = 0.009$, $p < 0.001$, respectively), but not worse than the Control ($p = 0.394$).

The interaction effect resulted from different learning rates between the groups (see Figure 14), where subjects learned in the following order (fastest to slowest): Concurrent Feedback, Terminal Feedback, Control, and Adaptive Threshold. The Concurrent Feedback group significantly outperformed the Control group for the first six Blocks. Unlike the Concurrent Feedback group, the Terminal Feedback and Adaptive Threshold groups started with the same initial performance as the Control group. The Terminal Feedback group learned more quickly than the Control group, and significantly outperformed the Control group for Blocks 4 and 5. Compared to the Control group, all groups performed at statistically similar level after Block 6. Investigating the immediate evaluation effects when the group-specific treatments are removed in Block 13, the mixed model shows no change in performance for any of the groups ($p > 0.989$ for all groups). As such, the percentage of successfully completed trials does not show any effect from the guidance hypothesis (i.e., the subjects did not rely on the feedback to complete the task and removing the feedback did not result in decreased performance).

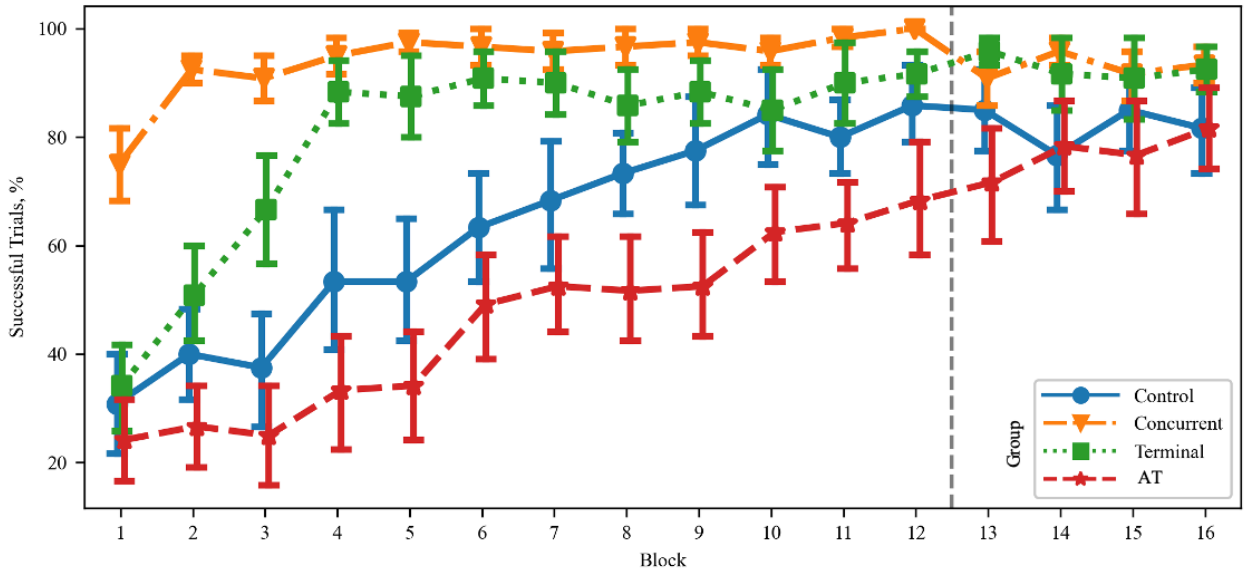



*Figure 14.* Percent of Successful Trials. The vertical dashed line represents the transition from training to evaluation. Error bars shown are the standard error of the mean. The Adaptive Threshold Group is abbreviated as "AT."

The TP was calculated for the evaluation phase and averaged across Blocks 13 through 16 (i.e., fourth Session). Throughput is generally used to assess an input device, which should be measured when

the subjects can complete the task. Since there were no significant differences in the evaluation phase for percent complete, we only calculated TP at this time as initial Blocks may have been biased towards low *IDs*. There was no significant difference in TP between the Groups ($F(3, 44) = 1.625, p = 0.197$). The mean TP for all subjects was found to be 0.564 ± 0.023 bits/s ($\mu \pm \sigma$).

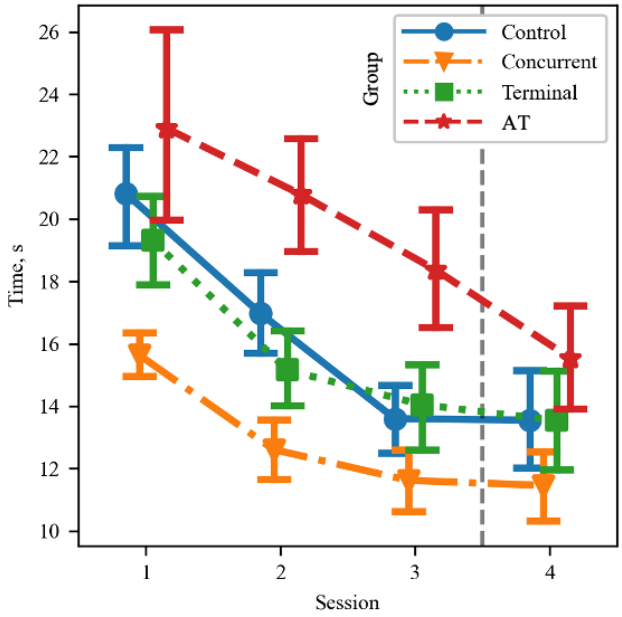



*Figure 15.* Completion Time by Session (set of 4 Blocks) across groups. The vertical dashed line represents the transition between the training and evaluation phase. The Adaptive Threshold Group is abbreviated as "AT."

The pseudo randomization of the 40 unique target positions occurred over four Blocks, thus it was appropriate to average completion time over a Session (i.e., set of four Blocks). Completion time was only defined for successfully completed trials; the results are displayed in Figure 15. There were significant main factors of Group ($F(3, 43.968) = 4.390, p = 0.009$) and Session ($F(3, 131.074) = 24.906, p < 0.001$). The interaction effect between Group and Session was not significant ($F(9, 131.069) = 0.782, p = 0.633$). A Tukey test showed that the Concurrent Feedback group performed significantly better than those in the Adaptive Threshold group ($p = 0.004$), which was the only significant difference between groups, and no groups significantly outperformed the Control group. Analysis of the Session factor

showed increased performance ($p < 0.05$) until the last two Sessions, which were not statistically different ($p = 0.646$). These results further supported that the guidance hypothesis did not occur.

The significant difference in completion time between the Concurrent Feedback and Adaptive Threshold groups may have been explained by increased nPL, but there were no significant group differences ($F(3, 44.099) = 0.908, p = 0.445$). Increased nPL correlated with higher *ID* targets (see Table 3).

*Table 3.* Normalized Path Efficiency for Different Targets

| Index of Difficulty (ID) | 1.00 | 1.58 | 2.32 | 3.17 | 4.09 |
|---|---|---|---|---|---|
| Average | 32.11 | 3.80 | 1.80 | 1.38 | 1.20 |
| Standard Error of the Mean | 2.90 | 0.29 | 0.08 | 0.04 | 0.02 |

### 3.5 Discussion

The study investigated the effects of automated training methodologies to efficiently and accurately use a sEMG command system, and elucidate the relationship between performance, workload, and trust during training. The command metrics from the Command Accuracy Test, which were administered before training (Test 1), after training (Test 2), and after evaluation (Test 3), largely did not indicate Group differences. We expected the Groups to perform similarly during Test 2, which was administered at the conclusion of the training phase and when the subjects should have reached a performance plateau at the end of training. However, in Test 3 we expected that the Adaptive Threshold group would optimize their inputs to the stabilized threshold, but we did not observe significant differences in Test 3. We did not expect to observe differences between Groups in Test 1, but there was a significant difference in average command amplitude between the Control and Adaptive Threshold and Concurrent Feedback and Adaptive Threshold. Prior to Test 1, the subjects across groups had received a generic experiment introduction, group-specific training tailored to their different interfaces, and a manual calibration of their signals. After investigating the data from Test 1, we concluded that the differences may be a result of the group specific training and the manual calibration. The manual calibration for the Adaptive Threshold group had to accommodate the

range of thresholds ($0.1 \leq l_1 \leq 0.3$), which increased the average normalization value by approximately 24% compared to the other groups. As a result, the processed signal input, $\bar{x}$, decreased by the same proportion for the same rms input value. Combined with the standard threshold value ($l_1 = 0.2$) during all Code Accuracy Tests, the Adaptive Group would have needed to input larger signals to compensate. There are no significant differences after Test 1, during the evaluation phase, or in the MVC (which would indicate fatigue), therefore the effect seems to be limited to pre-training. The Adaptive Threshold group's average command amplitude values did increase for Test 2 and Test 3 compared to Test 1. Tests 2 and 3 occurred after the Fitts's law-based task when the Adaptive Threshold group would have seen the effect of their input on the cursor motion. The Adaptive Threshold group did not pursue a strategy of maximally contracting their muscles evident by their convergent average command amplitude values with the other groups, particularly the Concurrent Feedback group which had feedback on threshold crossings. Our subjects' adaptation follows other studies that observed adaptation to input noise characteristics [103] and the refinement of muscle synergies [106]. The subjects in this study improved across the command metrics after the training phase, demonstrating adaptation to the sEMG command system. In terms of these command metrics, the specific automated training methodology was less important.

The Adaptive Threshold group had increased mapping uncertainty during training while the Concurrent Feedback and Terminal Feedback groups received augmented feedback about the sEMG control system. Augmented feedback was more effective in assisting subjects in achieving a performance plateau, requiring only a third of the prescribed training phase and following the findings of Basmajian [78]. The Control and Adaptive Threshold groups needed the entire training phase to achieve a similar performance level, indicating that the augmented feedback enhanced the early acquisition of motor learning skills. The early motor learning we observed, where the Concurrent group led the Terminal group in initial performance gains, was expected as this sEMG testbed was designed to investigate a challenging task. Functionally complex motor control tasks are generally expected to respond better to concurrent feedback [61], though it remains difficult to estimate a task's functional task complexity, or instances where Concurrent versus Terminal feedback will be more effective. Both feedback levels eventually resulted in

the same performance, but the high levels of performance initially observed in the Concurrent group (and corresponding decrease in cognitive workload), likely resulted from the subject's increased internal understanding of the task dynamics. The augmented feedback groups appear to have the strongest performance advantage.

However, the Adaptive Threshold group's results were interesting for two reasons. First, varying thresholds/increased mapping uncertainty did not appear to cause adverse effects compared to the Control group. Unreliable automation behavior can lead to poor human-automation interaction [107], but was not the case in this study. Secondly, the Adaptive Threshold group only trained at the standard threshold ($l_1 = 0.2$) for 20% of the trials compared to the other groups. The changing threshold did affect task metrics during the training phase; the varying threshold levels were not inconsequential as evident by the significant differences between the Adaptive Threshold and Concurrent Feedback groups for cognitive workload, trust, percent of successful trials, and completion time. However, the Adaptive Threshold group's performance was indistinguishable from the other groups during the evaluation. The training challenge of varying thresholds/increased mapping uncertainty did not cause poorer performance during evaluation, but also did not see significantly improved performance. It is possible that the adaptation training methodology could be redesigned to better elicit the benefits. Alternatively, we could have evaluated all groups at a novel threshold. To observe potential task generalization benefits, this methodology may be better tested with a different task that still uses the sEMG commands or a similar underlying task structure (a.k.a. a transfer task), instead of returning to a stable condition. For example, Braun et al. [82] showed that subjects trained with varying visuomotor tasks were able to quickly generalize to novel tasks that still retained similar features as the original training task. In one sense, the Adaptive Threshold group does have better command proficiency due to similar performance with less training at the standard threshold, which is encouraging although not strong support for the first hypothesis that the Adaptive Threshold group would have the highest command proficiency. We maintain that subsequent experiments should continue investigating the use of adaptation strategies for training.

The task results suggest that subjects reached a similar level of task proficiency by the end of the training phase, as we observed no significant differences between groups during the evaluation phase. However, a more complex interaction between task performance, cognitive workload, and trust occurred during the training phase and varied by training methodology. The Concurrent Feedback group started with the highest task performance, lowest cognitive workload, and highest level of trust. By the 5$^{th}$ Block, the Terminal Feedback group overlapped with the Concurrent Feedback group in terms of percent of successful trials, cognitive workload, and trust. These two groups then tracked each other for the remainder of the study. In both augmented feedback groups, our feedback displays act as an expert "Instructor" model, providing information to subjects in real-time about either the processed sEMG signal, $\bar{x}$, or their command state. Correspondingly, the results we observed closely align with prior literature when subjects are provided with expert feedback; subjects in both of our augmented feedback groups displayed reduced "cognitive load during initial practice, helping [them] integrate declarative knowledge with physical skills" [64]. The Concurrent Feedback and Terminal Feedback groups did not have significantly different average completion times throughout the study. In their initial learning, the Concurrent Feedback group demonstrated high initial performance with smaller, incremental gains, while the Terminal Feedback group had large Block-to-Block improvements. The Adaptive Threshold group did not appear to reach a performance plateau for the percent of successful trials and steadily, but not significantly, improved in subsequent evaluation Blocks. Interestingly, the trust score also continually increased for both the Control and Adaptive Threshold groups. Although the changing cursor dynamics in the Adaptive Threshold group does not seem to have adversely affected trust compared to the Control group, the Adaptive Threshold group did report a significantly higher cognitive workload for Blocks 9-11 than the Concurrent Feedback group. Once the threshold stabilized, the cognitive workload in the Adaptive Threshold group was not statistically different from the other groups. These effects that may be explained by Hoff and Bashir's three layer model [96], where the training methodology altered situational trust and the continued interaction affected learned trust. Jian et al.'s trust survey specifically asks subjects about their familiarity with and confidence in the system, and asks subjects to rate the reliability of the system and its outputs [104].

Changes in trust scores may reflect differences in components like confidence and reliability, but the survey also includes an explicit question about trust ("I can trust the system"). Analysis of the individual questions did not contradict the averaged results of the trust scale, however, and it likely that this reflects that learned trust, not simply understanding of the system, is increasing with time. Overall, these results suggested that augmented feedback led to earlier task performance gains with improvements in trust and cognitive workload. Surprisingly, the varying thresholds/increased mapping uncertainty did not adversely affect trust, but the mental cost was reflected in the task performance and cognitive workload. In general, the second, third, and fourth hypotheses from section *3.3.4 Hypotheses* were largely supported with increased trust and decreased workload by the evaluation phase, with no differences between groups.

The cognitive workload scores indicated that the task of learning a novel sEMG cursor control system to hit targets on a screen was challenging. The average cognitive workload scores by Block across Groups ranged from 6.71 to 3.65 with the highest average score occurring in the first Block. The average cognitive workload score decreased with subsequent Blocks. According to the Modified Bedford Workload Scale [89], a score of 7 indicated that the task was possible, but that there was "…minimal spare time for additional tasks." A score of 6 meant the task was possible and cognitive workload tolerable and "there was some but not enough spare time available for additional tasks" [89]. By the end of training, the average cognitive workload score was rated approximately at a 4, as the task was considered possible with tolerable workload and "…ample time to attend to additional tasks" [89]. In a simulated, four degree of freedom manually controlled spacecraft inspection task with a secondary task and verbal callouts, the average Modified Bedford Workload scores ranges from 7.17 to 3.94 [60]. The similarity in the workload scores indicate that the subjects found the task to learn sEMG control to be sufficiently difficult and enabled us to observe learning and performance improvements during training. We intended for the subjects to become proficient in sEMG cursor control by the end of training. Therefore, for an initial evaluation of automated training methodologies, we believe the task complexity to be sufficient. In future development of sEMG control for aerospace applications, it would be beneficial to include additional tasks to simulate a more realistic environment with representative aerospace tasks and displays.

There are not many active BoMIs for aerospace applications available for comparison and with similar metrics. The code accuracy reported in this study, 67.2% in Test 3, appeared similar to the classification accuracy of a BCI mouse at 73.1% [23]. The percent of successful trials during the evaluation phase (86.18%) exceeded that of a BCI mouse for a simplified, simulated spacecraft (66.7%) [22], and was similar to the success rate for crossing gates in gesture-based drone control (87.67%) [18]. However, the differences in the tasks and requirements for success limit the extent and interpretability of the comparisons. As another point of comparison, the throughput values during the evaluation phase for all groups fell within previously published results for sEMG cursor control systems. From *Chapter 2*, our previous single-site sEMG cursor control system with 2 DOFs (counterclockwise rotation and forward) reported 2.24 bits/s and 0.23 bits/s for control methodologies that used different levels of automation [1]. Multi-site sEMG systems have achieved 0.4 bits/s [49], 0.84 bits/s [51], and 1.3 bits/s [57]. The sEMG cursor control system used in this study had a throughput of 0.56 bits/s and may be of additional interest to the BCI community. The purpose of the sEMG cursor control system in this study was to extend the work presented in *Chapter 2* and to provide a testbed for automated training that lent itself to motor learning adaptation and was sufficiently challenging to probe the relationship between command and task performance, cognitive workload, and trust.

## 3.6 Conclusions

The study results largely supported our hypotheses described in section *3.3.4 Hypotheses*. The percent of successful trials performance during the training phase followed the order of Concurrent Feedback, Terminal Feedback, Control, and Adaptive Threshold until all groups reached a statistically similar performance level towards the end of the training phase. All groups performed similarly in percent of successful trials during the evaluation phase. The subjects' trust followed our expectations with Concurrent Feedback and Terminal Feedback having the highest levels, while the Control group's trust continually increased at a slow rate. The Adaptive Threshold group had lower trust during training, and the trust increased to the level of the other groups during evaluation.

The cognitive workload results also supported our hypotheses that Concurrent Feedback and Terminal Feedback groups would have the largest decrease in cognitive workload, and that all groups would have similar cognitive workload during the evaluation phase. Interestingly, the Concurrent Feedback and Terminal Feedback groups converged across performance, cognitive workload, and trust by Block 5. Our hypothesis that the Adaptive Threshold group would achieve the highest command proficiency was not entirely supported, although we observed encouraging indications. In future studies it may be of interest to include a transfer task or conduct the Command Accuracy Test throughout the training phase to improve understanding of the adaptation methodology. This study provided insights on the relationship between performance, cognitive workload, and trust for various automated training methodologies, and highlighted the advantage of augmented feedback methodologies during the training phase. By directly comparing methodologies and identifying human-automation interaction effects, this research provides a quantitative basis for exploring combined training approaches in future research. It would be interesting to investigate a mixed training methodology approach that combines concurrent feedback with the motor adaptation approach, where subjects potentially benefit from both increased feedback, adaptation rate, and skill generalization. This research also contributes to training users of BoMIs, but it is not known if the study results generalize to BoMIs with other signal types. The muscle activation patterns were mapped to cursor commands, which were relevant to the task, but could be translated to other commands as needed for an application and sEMG control is not limited to an on-screen application. The availability of multiple communication modalities provides flexibility and the opportunity to create new interaction paradigms.

This study addressed outstanding questions from *Chapter 2*. The changes made by decreasing the timeout between inputs and reduced trial time did not appear to adversely impact performance. Increasing the number of trials from 80 to 160 appeared adequate for training all groups to a similar level of proficiency in the evaluation phase. There were no significant differences in the command metrics between after training and after evaluation, further supporting that the training period of 120 was appropriate. This study also confirmed that the concurrent feedback, which was used in *Chapter 2*, resulted in early, high-level performance. The Concurrent Feedback group plateaued after Block 5, or 50 trials, for percent success and

Block 12, or 120 trials, for completion time. (The completion time was averaged over sets of four Blocks; it is possible the performance plateau for completion time may occur earlier.) This study demonstrated that concurrent feedback quickly trained subjects to use a myoelectric control system with an increased number of commands. These results built confidence in the abilities of subjects to learn less intuitive, serial patterns for myoelectric control, which expands the possibilities for single-site sEMG.

# Chapter 4

# Myoelectric Command Accuracy Improvement in Trained Users

## Preface

The Coronavirus pandemic disrupted the originally planned research due to limitations on in-person research activities, which was needed for myoelectric control studies. The restrictions on in-person research provided an opportunity to revisit the dataset from *Chapter 3* and further investigate myoelectric command accuracy performance. *Chapter* 3 explored questions concerning feedback and training, but did not investigate predetermined timing requirements, such as the timeout between inputs and the input durations. The purpose of the study presented here in *Chapter 4* was to determine if the "peak" performance achieved by the end of the original study could theoretically be improved with the addition of classifiers and/or adaptable features to alter the timing requirements. This work was originally published in [4].

## 4.1 Introduction

Myoelectric control, or the use of EMG signals to convey a command, has been used in prosthetics and HCIs, with various strategies employed to create a reliable communication interface. However, EMG can suffer from signal non-stationarities due to subject fatigue, shifting electrodes, and postural changes [28]. Some researchers have mitigated signal non-stationarities by developing adaptive systems. These adaptive systems have been shown to improve pattern recognition classifier accuracy during increased signal noise [108], to augment performance with online co-adaptive learning [109], [110], and to maintain repeatability over multiple sessions [111]. All of these research studies used sEMG control with multiple sensors on the forearm, and most used pattern recognition [108] or regression-based methods [109], [110].

Adaptive sEMG systems can focus on mitigating signal non-stationarities and/or improving overall system accuracy and performance. Benchabane et al. [111] developed a system for simultaneous control of prosthetic fingers using an adaptive threshold to select one or more finger functions and demonstrated low

accuracy loss, despite multiple sessions (-4.8% change in accuracy) and subject fatigue (-2.4% change in accuracy). Similarly, Zhang et al. [108] tested their system with several noise levels to compare a Linear Discriminant Analysis (LDA) classifier with an adaptive LDA (ALDA) that periodically incorporated new training data to recalculate parameters online. The ALDA outperformed the LDA across conditions [108]. Hahne et al. [109] investigated different adaptation rates for a co-adaptation cursor-to-target task in a regression-based system to improve task performance. In a similar regression-based system, Yeung et al. [110] compared different methods of weighting new training data, tested several adaptation rates, and demonstrated improved performance compared to an unassisted user. Taken together, these selected studies demonstrate the promise of adaptive sEMG systems for novice users. Our work was motivated to investigate the benefits of adaptive systems for users after completing training and reaching a learning plateau to further enhance their performance.

In *Chapter 3*, we developed a single-site sEMG cursor control method that allowed the user to select among four cursor commands, using a coded sequence of short and long muscle activations. This method may be less sensitive to signal non-stationarities due to the system design emphasis on user training and adaptation as opposed to the machine. The user only needs to cross a threshold to communicate and can adjust their own actions to make corrections when the signal changes. However, this approach placed the burden of learning on user training, therefore we investigated the effects of various training strategies on performance, cognitive workload, and trust [3]. At the end of the training period, subjects across training strategies successfully completed 86% of trials and had a command classification accuracy of 66% for a cursor-to-target task [3]. The purpose of the work presented in this chapter was to understand if adding an adaptive timing component or an individually trained classifier would improve the classification accuracy of trained subjects.

## 4.2 Methods

### *4.2.1 Experiment Design and Setup*

In this study, we utilized the data collected in [3] to explore the possibility of integrating an adaptive feature or classifier to our sEMG control method. (The detailed experiment design and protocol can be found in *Chapter 3.*) The forty-eight subjects that completed the experiment ranged in age from 18 to 23 years with two left-hand dominant subjects and an equal participation of men and women. Subjects were able-bodied university students who reported no history of neuromuscular disorders or myoelectric control experience. The UC Davis Institutional Review Board approved the experimental protocol.

The experimental setup included two electrodes (ConMed 1620 Ag/AgCl center snap) placed approximately 2.5 cm apart near the extensor digitorum proximal attachment on the dominant side. The reference electrode was fixed close to the lateral epicondyle of the humerus (see Figure 5b). The signal acquisition followed [54] and was processed as in [1], where the signal was sampled at 4096 Hz for 256-sample windows with a Butterworth filter ($4^{th}$ order, bandpass at 10 Hz and 500 Hz). The rms value for the window was normalized by each subject's calibration value and put in a moving average filter (length 0.5 s) for a processed signal update rate of 16 Hz.

### *4.2.2 Cursor Control and Training Methods*

As described in *Chapter 3*, The sEMG method consisted of four cursor-control commands—up, down, left, right—that the subjects selected through a serial pattern akin to Morse code [53]. The commands consisted of two consecutive signal inputs with a rest between each input; if the rest exceeded the timeout (0.5 s), the pattern reset (see Figure 16A). An input started when the processed signal exceeded a threshold, $l_1$, and ended when the processed signal returned to below the threshold. Inputs less than 0.5 s were categorized as "short" and inputs greater than 0.5 s were "long". The combined pair of short and long inputs (a.k.a. muscle activations) defined a command; for example, two short inputs selected the "up" command (see Figure 16B). A third input moved the cursor forward in the selected direction while the processed signal remained above the threshold.

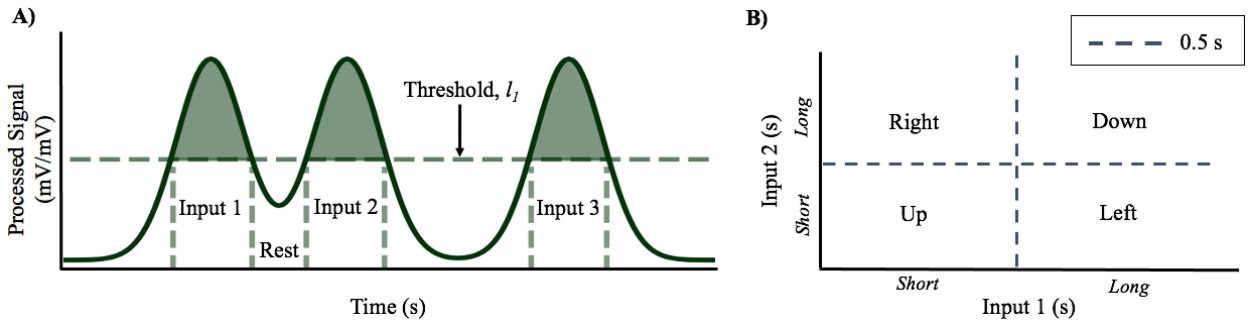



*Figure 16.* Overview of the coded sequence used in the sEMG method (also shown in *Figure 7A*). A) An illustration of a processed EMG signal is shown here. The coded sequence requires two inputs above a threshold to select the command and the third input results in continuous, forward motion until the signal drops below the threshold. Subjects must rest (i.e., drop below the threshold) between inputs, but a rest greater than 0.5 s resets the sequence. B) The commands are defined by the combination of short (≤ 0.5 s) and long (> 0.5 s) inputs. For example, the "down" command is two long inputs.

During the training portion of the experiment, the subjects learned to control a cursor on a computer screen to select targets in a Fitts's law-based task. There were 120 training trials and 40 evaluation trials, where a trial was a single cursor-to-target attempt. The training strategy was in effect during the training trials and removed during evaluation. Subjects were assigned to one of four groups that determined the training strategy: Repetition[2], Concurrent Feedback, Terminal Feedback, and Adaptive Threshold.

1) The Repetition group learned only by repeatedly performing the cursor-to-target task.

2) The Concurrent Feedback group received an additional visual feedback element, where the cursor body changed color when the processed signal exceeded the threshold shown in Figure 16A (also may refer to Figure 7 for feedback timing).

3) The Terminal Feedback group's additional visual feedback element indicated the selected command at the onset of Input 3 (see Figure 16A; may refer to Figure 7 for feedback timing).

[2] The groups in *Chapter 4* include a Control Condition with the original timing requirements to compare with the adaptive features and classifier groups. For clarity the Control Group from the original experiment covered in *Chapter 3* is referred to as the Repetition group here. Both Condition and Group are factors in the statistical analysis.

4) The Adaptive Threshold group saw the same interface as the Repetition group, but the threshold varied trial-by-trial ($l_1$ = 0.10, 0.15, 0.20, 0.25, 0.30). The threshold was set to 0.20 for other groups (threshold shown in Figure 16A).

### *4.2.3 Analysis*

To assess the subjects' proficiencies in selecting commands, a Command Accuracy Test was administered before training, after training (Test 2), and after evaluation (Test 3). Subjects were prompted to produce each command 5 times for a total of 20 attempts per test. Subjects had only one try within an attempt. Command classification accuracy was calculated as the number of successfully selected commands out of the total for each test. From the previous study, there were no significant differences after the training period between Tests 2 and 3 [3]. An independent, two-tailed t-test indicated that there were no significant differences in command accuracy between men and women ($p > 0.7$). Therefore, the data from the Command Accuracy Tests 2 and 3 were combined and used for the analysis presented here. A larger dataset would have been ideal, but this analysis was developed under research limitations imposed by the Coronavirus pandemic, and so only our previously existing data from [3] was used.

In the original study, successfully selected commands met the input duration requirements (short ≤ 0.5 s, long > 0.5 s) and the rest between inputs did not exceed a timeout (> 0.5 s). These values were selected from pilot studies and may not be optimal in general or for specific individuals. For the purposes of this study, the classification accuracy calculated with these predetermined values was considered the Control. We then recalculated the classification accuracy for two additional conditions that included certain unsuccessful commands: 1) No Timeout and 2) Individual Classification. For the No Timeout condition, the first two inputs were classified according to the original input duration requirements, but without regard to the timeout (rest > 0.5 s, rest shown in Figure 16A). The Individual Classification used the first two inputs as features for supervised machine learning and did not exclude any observations that exceeded the original timing requirements. Table 4 summarizes the Conditions with respect to the timing requirements used in this study.

*Table 4.* Timing Requirements per Condition

| Condition | Timeout | Input Duration |
|---|---|---|
| Control | Yes | Yes |
| No Timeout | No | Yes |
| Individual Classification | No | No |

In the Individual Classification condition, we made several decisions to accommodate the limited number of observations per subject. The K Nearest Neighbors (KNN) algorithm was selected due to its straightforward implementation without the need to tune many parameters. The Python package *scikit-learn* [112] was used and the parameters for $k$ and *weights* were set to 4 and "distance" (voting weighted by distance from the observation to the $k^{th}$ neighbor), respectively. The number of neighbors, $k$, was selected based on the number of observations per class and approximately followed the general rule for selection of $k$ ($\sqrt{observations} \approx$ 5-6). We assumed that subjects performed command selection with precision, therefore we assigned more weight to closer neighbors. Typically, a 5- or 10-fold cross-validation strategy is used but based on the number of observations per class and the decision to use stratified folds, we implemented a 4-fold cross-validation with 3 iterations. Stratified folds aim to have even representation, or the number of observations, in each class. There were four classes, one for each command: up, down, left, right. Each fold was still unique, and we used a train/test split of 80/20. The number of observations in the training set ranged from 29-32, but we held the testing set constant at 8 observations (2 per class). One subject had a total of 24 useable observations, which resulted in a 2-fold cross-validation with 6 iterations and 4 neighbors. Unusable observations occurred when the subjects' attempts contained one or no inputs and accounted for 1.51% of the total attempts.

We compared the average classification accuracy across the Control, No Timeout, and Individual Classification conditions and hypothesized that removing the predetermined timing requirements would yield better classification accuracy. We also calculated the average increase in accuracy from the Control to No Timeout conditions, and the No Timeout to Individual Classification conditions. The reserved test

dataset was used for another Individual Classification accuracy estimate. Results are reported as ($\mu \pm \sigma$) unless shown otherwise.

### 4.3 Results

We ran a two-factor mixed model with a between-subjects factor of Group and a within-subjects factor of Condition. Significant effects ($p < 0.05$) were analyzed with the Tukey HSD test with the Satterthwaite method to calculate the degrees of freedom. The main effect of Condition was significant ($F(2,88) = 31.40$, $p < 0.0001$), but not Group ($F(3,44) = 1.86$, $p = 0.15$) nor the interaction between Group and Condition ($F(6,88) = 0.83$, $p = 0.55$). The Tukey test for Condition showed significant differences for each pairwise comparison. The classification accuracy improved from the Control to No Timeout ($p < 0.0001$), Control to Individual Classification ($p < 0.0001$), and from the No Timeout to Individual Classification ($p = 0.04$). The Conditions achieved average classifications accuracies of 66.15 ± 26.31% for the Control, 76.56 ± 21.52% for No Timeout, and 81.58 ± 20.30% for Individual Classification; the results are shown in Figure 17.

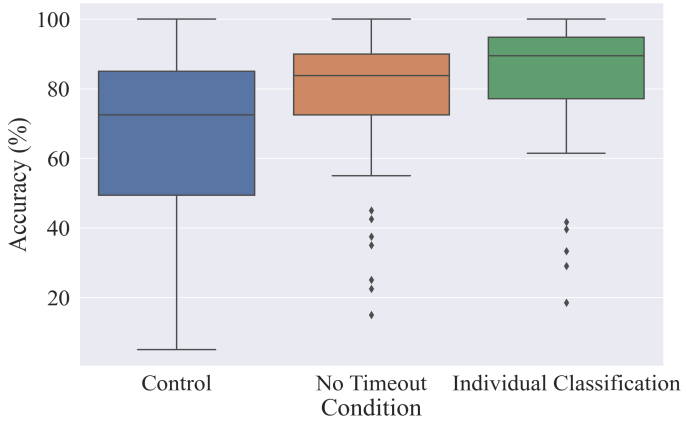



*Figure 17.* Classification Accuracy across Conditions. The box-and-whiskers plot displays the median value and 1st and 3rd quartile within the box. The whiskers extend up to 1.5*x* of the 1st and 3rd quartile (not exceeding the maximum and minimum data points). Data points outside the whiskers are depicted as markers.

Classification accuracy results varied as some subjects scored 100% accuracy in the Control and did not benefit from the other Conditions, whereas other subjects had large performance gains with loosened timing requirements. For the No Timeout condition, the increase in percentage points ranged from 0.00 to 50.00 points with an average increase of 10.42 ± 13.71 points. The Individual Classification condition inherently included the No Timeout consideration because both did not impose the timeout restriction (see Table 4). The change in classification accuracy percentage points from No Timeout to Individual Classification conditions ranged from -12.29 to 36.46 points with an average increase of 5.02 ± 10.29 points. For the cases where the classification accuracies decreased from No Timeout condition ($N$ = 17), the Individual Classification condition classification accuracies were either worse ($N$ = 8), better ($N$ = 8), or the same ($N$ = 1) as compared to the Control condition.

The reserved test dataset also provided another classification accuracy estimate for the Individual Classification condition. The resulting accuracy across groups was 85.00 ± 18.36%. These results were similar to the classification accuracy estimated from the training dataset with cross-validation. The classification accuracy from the training dataset was used in the statistical analysis because it was estimated over more observations and may be more accurate.

### 4.4 Discussion

The results supported our hypothesis that removing predetermined timing requirements would yield better classification accuracy and indicated that most subjects could potentially benefit from customized adaptive elements in our sEMG command system. The main effect of Group was found to be not significant; therefore the training history did not appear to be a factor in the outcome of the conditions, and was consistent with our prior results of no significant performance differences after training [3].

The No Timeout condition either benefited ($N$ = 33) or did not affect ($N$ = 15) the classification accuracy. Within the combined Test 2 and 3 datasets across groups, the timeouts accounted for 16.98% of the total attempts. It was interesting to note that timeouts by Group occurred in 15.63%, 5.42%, 17.92%,

and 28.96% of the attempts for the Repetition, Concurrent Feedback, Terminal Feedback, and Adaptive Threshold groups, respectively. The Concurrent Feedback group may have been more attuned to the timeout requirement because the visual aid inherently gave information about the timing between inputs. The Adaptive Threshold group may have struggled learning the timing with the changing threshold. The timeouts had a mean of 0.41 s and median of 0.31 s. For an online implementation, the timeout could be modified to the subject's maximum timeout or 95$^{th}$ percentile based on a Command Accuracy Test administered after a training period.

The Individual Classification improved the estimated classification accuracy for most subjects. Figures Figure 18 through Figure 20 display the input durations performed for the target commands, which have a similar layout to Figure 16B with lines indicating the predetermined input duration. The subject's data shown in Figure 18 demonstrates the benefit of Individual Classification when command inputs are precise but stray from the prescribed timing. In this case, the subject saw no improvement from the Control condition to the No Timeout condition which both had classification accuracies of 95%. Instead, the subject benefited in the Individual Classification condition with flexible input durations and improved to a 100% classification accuracy. In contrast, some subjects' estimated classification accuracies remained low or decreased in the Individual Classification condition. One possible explanation may be that lower performing subjects inconsistently produced inputs, such that the classes are not easily separable, and the nearest neighbors mostly belong to other classes. This case is shown in Figure 19 and there is limited separation or clustering. Alternatively, high performers may have produced inputs within the prescribed timing, but the observation occurred closer to other classes. In Figure 20, the subject performed with 100% classification accuracy in the Control condition, was unaffected by the No Timeout condition, and had a decrease in classification accuracy to 97%. The encircled observation fell within the correct quadrant of the plot, but too close to neighbors of a different class causing the misclassification.

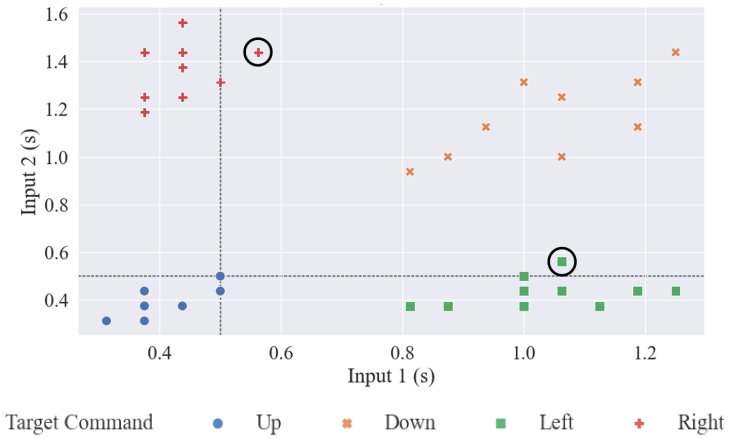



*Figure 18.* Example of a higher performing subject with classification accuracies of 95% (Control), 95% (No Timeout), and 100% (Individual Classification). The encircled points show attempts that did not have the correct input duration but remained close to their classes.

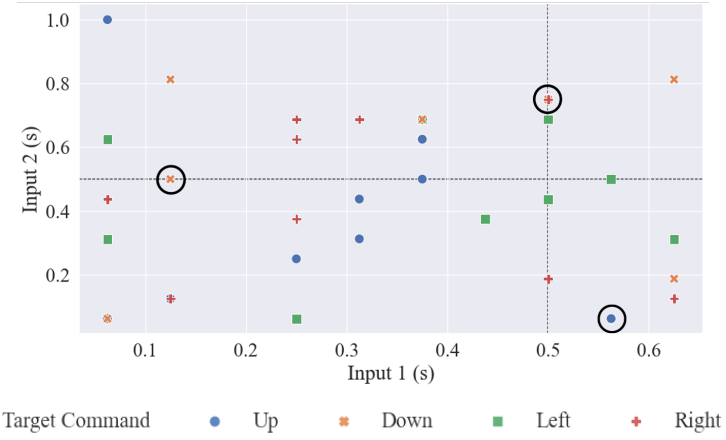



*Figure 19.* Example of a lower performing subject with classification accuracies of 30% (Control), 35% (No Timeout), and 41.67% (Individual Classification). The encircled points show attempts that did not have the correct input duration and/or were surrounded by other classes.

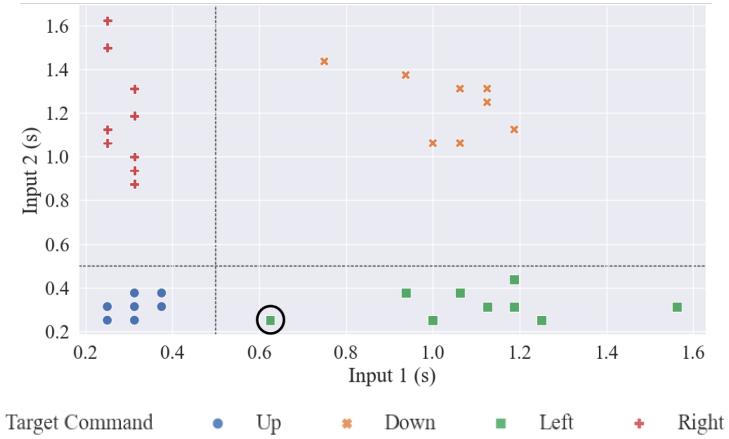



*Figure 20.* Example of a higher performing subject with classification accuracies of 100% (Control), 100% (No Timeout), and 96.88% (Individual Classification). The encircled point had the correct input durations but was closer in distance to the Up class causing a misclassification.

In our study, the original classification accuracy ranged from 5.00% to 100.00% with low and high performers within the subjects. Classifiers and adaptive features may improve the average performance but should target individuals that would benefit from customized timing requirements. For example, the subject whose data were depicted in Figure 18 did not benefit from removing the timeout restriction but improved with Individual Classification. For Individual Classification, the average classification accuracy of 81.58% was comparable to Benchabane et al. [111] results of 87.8% average accuracy. Also, eleven of our subjects achieved classification accuracies of greater than 95%, similar to the LDA average accuracy of 95.12% and the ALDA classifiers with 95.45% to 100.00% accuracy [108]. Our results suggested that a customized approach to implementing classifiers and adaptive features could improve performance of already trained subjects. This customization could occur after a training period using the data from a Command Accuracy Test to determine the best approach for the subsequent online implementation.

This work used data already collected from subjects in an experiment not designed to test this hypothesis. The two main limitations of this study were the small dataset and post-data collection analysis. The maximum of 40 total observations was smaller than suggested by other researchers, who recommend starting with 75 observations per class [113]. The limited dataset for individuals also prevented a robust model selection step, but analysis done at the group level indicated a preference for KNN and selected parameters. Overall, the classification accuracy estimates indicated the potential benefits for individual subjects.

These are *potential* benefits. It is unclear whether an online adaptive sEMG command system that allows for custom timing would yield improved accuracy. How the subject would adapt in real-time is also unknown. Another possibility would be to use the adaptive component to train the subjects to become more efficient (e.g., smaller command time). Overall, these results are encouraging to pursue a future study to address these questions.

## 4.5 Conclusions

The analysis conducted in the study presented in this chapter indicated that *potential* performance improvements for fully trained subjects may be possible by incorporating adaptive features and/or classifiers for customized timing. In a future validation study, the customized timing should take into consideration the individual's performance. Since the analysis was conducted with previously collected data and it was not possible to test the conditions online, the realized benefits remain unclear. Furthermore, it was not evident that the timing requirements should change for the training period. In both *Chapter 2* and *Chapter 3*, the subjects learned to vary sEMG input durations to select commands and perform the task to a high degree of success by the end of the experiment. Therefore, it appears reasonable to continue to use 0.5 s to separate short and long inputs.

A factor that can limit the efficiency of a sEMG command system based on a serial pattern of short and long inputs is the minimum command time. Any commands that include a long input require an input of at least 0.5 s in duration to communicate "long." After an initial training period, it may be beneficial to use an adaptive feature to coach the subjects towards shorter inputs, but with enough precision that short and long inputs are clearly distinct. Based on the findings of *Chapters 3* and *4*, pairing an adaptive feature with concurrent bandwidth feedback could result in decreased command time. It would be interesting to study the upper communication rate limits in subjects.

# Chapter 5

# Myoelectric Control of Robot During Coordination Task

## Preface

The goal of the study conducted in this chapter was to perform an in-person, pilot study for a coordination task with myoelectric control of a collaborative robot that could function as a supernumerary robot. The coronavirus pandemic impacted the study schedule, such that I conducted a pilot study with a limited number of subjects and reduced the in-person duration of the study. I used the data from *Chapters 2* and *3* to estimate the number of trials, employed concurrent feedback for efficient training, and developed a coordination task that used the subject's two natural hands and a third robotic "hand" controlled via sEMG on the subject's leg.

## 5.1 Introduction

Robots are used in a wide variety of applications including disaster response and minimally invasive surgery. In these cases, robots augment the capability of humans by extending their presence into an extreme or challenging environment. The research interest in human-robot collaboration has grown exponentially [114] and a related field has recently emerged: supernumerary robots (SRs). SRs are characterized by creating an additional, kinematically independent limb or appendage that may serve various functions. Of particular interest is SRs' potential to reduce risk and increase capability in challenging environments by providing a third, robotic arm for manipulation. The current state-of-the-art in SRs has focused on the development of the robot and less research has been done on the user interface and performance of the human-SR system. The experiments conducted with humans and SRs to complete a task have focused on specific use cases and tasks that do not provide as much detail on the human-robot interaction.

SRs are distinct from other assistive robots and comprise a relatively novel field of research. Exoskeletons augment the abilities of existing limbs, prosthetics restore functionality by replacing lost

appendages, and supernumerary robots can create limbs that are kinematically independent from the user. SRs may manifest as arms [115]–[120], legs [121], [122], or fingers [33], [123]–[127]. Current supernumerary robots have several potential use cases: body bracing [115], [121], [122], overhead tasks [116], [117], grasp assistance [123]–[125], and manipulation tasks [118]–[120] (a more detailed proposed usage taxonomy of a wearable supernumerary robotic arm may be found in Ref. [128]). The research in this field has centered on device development and proof-of-concept but has generally not focused on the human interface and human-robot performance.

According to the classification framework proposed by Leigh et al. [129], human-robot systems that can be considered a single unit, may be classified by type of support and control methods. SR development has focused on indirect control, or "Pseudo-Mapping" and "Shared Control" in Leigh's framework, by predicting the human's intent (for examples see Refs. [116], [117], [123]). In contrast to using prediction to identify intent, there has also been some work on direct control of SRs (i.e., "Direct Control"). A supernumerary finger used a combination of hand gestures to control the position and sEMG to modulate the grip strength [33], and a subsequent version of the supernumerary finger incorporated buttons and haptic feedback [124]. Another SR, a supernumerary arm, was commanded by foot position and toe flexion to directly control 6-DOFs [119]. The selection of the control method, whether direct or indirect, may depend on the application including the role of the SR in the task. However, there has been little available guidance or knowledge regarding control methods and task allocation due to the novelty of the field. Therefore, this study aimed to focus on direct control for a primary task, which has not been well-studied in the field.

Direct control SRs that use BoMIs need to use an existing part of the body, which may seem counter to the concept of SRs creating a limb. However, for tasks or applications where part of the body would not otherwise be used, it provides an opportunity to reallocate resources. For example, the lower limbs are available in seated tasks as demonstrated in [119] for the foot control of a 6-DOF SR. Furthermore, coordination between the hands and legs occurs in a variety of skilled tasks, such as the use of foot pedals by dentists and doctors. For example, in a virtual laparoscopic task at least one hand worked in coordination

with the foot for 50% of the task time [130]. Alternatively, instead of controlling an instrument, the foot could also be represented as a hand. Abdi et al. compared two-handed and three-handed performance during a virtual task where subjects needed to catch falling objects [131]. The subjects controlled the third hand with their leg. The results showed that subjects missed fewer objects in the three-hand paradigm as task difficulty increased [131]. A similar study assessed coordination between the hands and legs for various virtual tasks and deemed the control strategy feasible for SRs [132]. However, the experiment did not include any robot simulation or hardware. Overall, the incorporation of foot control for SRs has precedent from other fields (refer to Ref. [133] for review of foot control in HCIs).

The purpose of this pilot study was to incorporate a direct control BoMI with a SR and assess the human's ability to coordinate between their hands and the lower limb while controlling the SR (the robotic, third hand). The task was designed such that all three hands were needed to successfully complete the task. Simultaneous coordination of the hands was the goal communicated to the subjects, but it was possible to operate each hand independently. Instead of designing our own SR, we used a commercial, collaborative robot. We also did not constrain the SR to be physically attached to the human, since a wearable SR may not be necessary for every application. This study addressed current gaps in direct control interfaces and limb coordination for SR applications.

## 5.2 Materials and Methods

### *5.2.1 Subject Recruitment and Demographics*

We designed a pilot experiment for preliminary investigation of the three-handed, coordination task with two natural hands and a robotic, third one. The protocol was approved by UC Davis Institutional Review Board, and adult participants were recruited from the university. Exclusion criteria included history of neurological/neuromuscular disorders, limitations on arm and leg mobility, and failing coronavirus-specific screening. We chose to limit the age range of the participants to 18 through 39 years old based on the risk rates published by the Centers for Disease Control and Prevention [134]. Participants had to show proof of

a negative coronavirus test and pass a coronavirus survey addressing symptoms and exposure risk on the day of their scheduled experiment session.

To further reduce the risk of coronavirus transmission, in addition to the coronavirus-specific screening, we remotely conducted the intake. At the start of the scheduled experiment session, the subject called the researcher to complete the coronavirus-specific screening. After passing the screening, the researcher reviewed the consent form with the subject; both the subject and researcher could view and sign the form online through DocuSign® [135]. Consented subjects then completed a pre-session survey to collect demographic information. The remainder of the experiment protocol occurred in-person at the research location. The in-person protocol consisted of experiment instructions, sEMG setup and calibration, a maximum voluntary contraction measurement, the task, a second maximum voluntary contraction measurement, and a post-session survey.

Four subjects consented to participate in this pilot experiment. The subjects ranged in age from 20 to 25 years with an equal participation of females and males. One subject identified their ethnicity as Hispanic, Latino/a, or Spanish and race as Asian. The remaining subjects were not of Hispanic, Latino/a, or Spanish ethnicity and identified their race as White. Origin and race survey questions followed the guidance document provided by the Food and Drug Administration [136]. All subjects self-reported a right hand dominance, and leg preference was determined using the revised Waterloo Footedness Questionnaire (WFQ-R) [137]. The leg preference was used for sEMG electrode placement, and three subjects reported a right leg preference. The remaining subject indicated a similar preference between left and right, but a recent injury resulted in the selection of the left leg. The subjects had some prior experience that may have made them amenable to the experiment's task. One subject had prior experience with myoelectric control, and all subjects played video games. Two subjects had significant prior experience with robots and the other two subjects had little to none, but all subjects answered that they thought they would be comfortable around robots.

### *5.2.2 Experiment Setup and System Architecture*

The pilot study aimed to observe coordination between two natural hands and a robotic, third hand controlled by sEMG from the leg. As this was our first experiment with this robot (UR5e, Universal Robots [138]), we decided to not allow direct, physical interaction. A robotic simulation task was originally considered but due to schedule constraints caused by the coronavirus pandemic, we opted to focus our efforts on myoelectric control with the physical hardware. Therefore, we used a computer-based task for indirect, physical interaction and direct myoelectric control and integration with the actual hardware. The overall experiment task software framework was created using AxoPy [99], which provides basic infrastructure to design and run myoelectric control experiments. Communication with the robot controller used the Universal Robots Real-Time Date Exchange (RTDE) interface [139] and leveraged the API created by the University of Southern Denmark (SDU) Robotics [140].

The subjects completed a cursor-to-target task visualized on a desktop monitor. They used their natural hands to input keyboard commands to control the cursor in 2-DOFs and used their leg muscle to command the robot end effector position in the third DOF. The cursor position on the screen in the third DOF reflected the actual, scaled position of the robot's end effector. The ConMed 1620 Ag/AgCl center snap electrodes were placed approximately 2.5 cm apart on the tibialis anterior below the lateral tibial condyle to measure the muscle activation and a reference electrode was affixed to the kneecap (see Figure 21). The electrodes were placed on the preferred leg as indicated by the WFQ-R. As in the previous chapters, the sEMG signal acquisition followed [54] and was processed as described in [1], where the signal was sampled at 4096 Hz for 256-sample windows with a $4^{th}$ order Butterworth filter (bandpass at 10 Hz and 500 Hz). The rms value for the window was normalized by each subject's calibration value and put in a moving average filter (length 0.5 s) to yield a processed signal update rate of 16 Hz. Commands and position data were exchanged between the experiment computer and robot controller at 16 Hz.

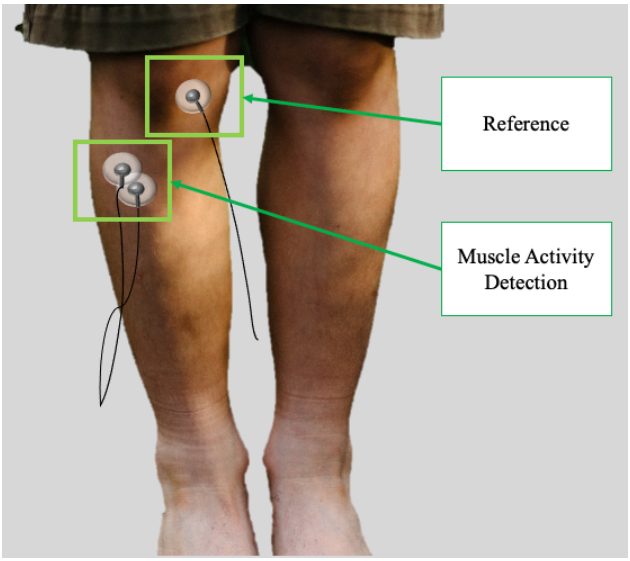



*Figure 21.* Electrode Placement on Tibialis Anterior. Image modified from [141].

During the in-person portion of the experiment, the subjects sat at a desk in front of a desktop monitor. Subjects were instructed to adjust the chair height, so their feet rested comfortably on the floor and their thighs were approximately parallel to the floor. The robot was positioned within the field of view and off to the right of the desktop monitor. The subject and robot were separated by an adequate distance and stanchions were used as a physical barrier marking the keep out zone. The layout of the experiment room is shown in Figure 22.

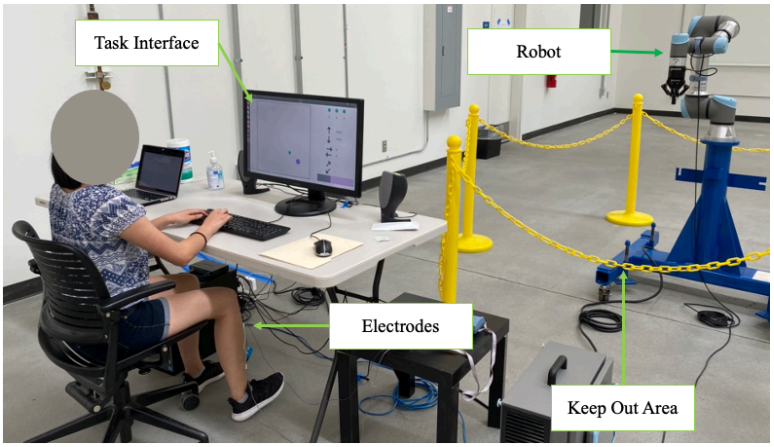



*Figure 22.* Experiment Room Layout.

### *5.2.3 Coordination Task*

The subjects were informed during the experiment instructions that their task was to "move the cursor in three dimensions to select targets on a screen using three 'hands'" with the goal of simultaneous movement in 3-DOFs. The subjects controlled the cursor motion using keyboard inputs with their hands and myoelectric control with their leg. Their left hands used keys *a* and *d* to move the cursor left and right along the x-axis, respectively. Each keypress increased or decreased the cursor speed in the x-direction meaning that a single press of key *a* would cause the cursor to continue moving left until either the cursor hit the task boundary and stopped, or a press of key *d* negated the leftward motion. The right hand manipulated the cursor motion along the z-axis, which was into and out of the screen. To show 3D motion in a 2D visualization, the cursor diameter decreased to show motion into the screen and the diameter increased for motion out of the screen. Key *j* decreased cursor diameter and key *l* increased the diameter. The cursor diameter was constrained within $0.15x$ and $2.45x$ of the original diameter. The subjects' legs controlled robotic motion in the y-axis to move the cursor vertically. Similar to *Chapter 2*, the subjects needed two commands, up and down, and used two inputs. The first input selected the command and the second input resulted in forward motion for the duration that the processed sEMG signal remained above the threshold, $l_1$. A short input ($\leq 0.5\ s$) selected the "up" command and a long input ($> 0.5\ s$) selected the "down" command. Unlike our previous work [1], [2], the forward velocity was not proportional to the processed sEMG value, $\bar{x}$. The robot moved at a constant velocity and therefore so did the cursor in the y-axis. This decision was made to help ensure smooth robotic motion. The commands are summarized in Table 5.

The user interface included an information interface and a task interface for the cursor-to-target task (see Figure 23). The information interface contained a set of status lights, the command key, and an sEMG signal bar. The status lights indicated in which DOFs the cursor was currently in motion by turning from gray to green. These lights could be helpful when the cursor moved at a low velocity, especially in the *z* axis, and provided a different visual representation of the subjects' goal, 3-DOF movement (i.e., keep all the status lights green). The subjects never had to memorize the cursor commands and could refer to the

command key. The sEMG signal bar displayed their current processed signal value as an overlaid dark gray bar. The light blue region and light pink region designated the rest area and active area, respectively.

*Table 5.* Summary of 3-DOF Cursor Commands

| Axis | Direction | Command | Input Method | Leg/Hand |
|---|---|---|---|---|
| *x* | left | *a* | keyboard | left hand |
| | right | *d* | | |
| *y* | up | *short* | sEMG | leg |
| | down | *long* | | |
| *z* | in | *j* | keyboard | right hand |
| | out | *l* | | |

In addition to the elements in the information interface, we used color feedback to provide additional information during the task. The cursor body had concurrent feedback to indicate the reception and interpretation of the sEMG signal (see Figure 3). The cursor body changed from black to a dark gold when the sEMG signal crossed the threshold; the color would change to light blue if the first input time exceeded 0.5 s to indicate a long input. The cursor body remained the color of the selected command during the second input. This concurrent color feedback was the same scheme as the manual rotate method in *Chapter 2*. The other two forms of color feedback applied to the target. In order to select the target, the cursor needed to be at approximately the same depth, which was represented by diameter. The cursor was considered at the same depth if difference in scale between the cursor and target diameters was within 0.10 units. In preliminary testing, we tried a fixed percentage of the target diameter, but the margins were too small for far targets and too large for near targets. Since it could be visually difficult to estimate the cursor depth, we used color feedback on the target body to indicate when the cursor and target were at approximately the same depth. The target was nominally light purple and turned green when the cursor was at depth. In addition to placing the cursor at the same depth as the target, the subjects needed to dwell on the target for 1 s to select it. The target turned orange when the cursor selected the target. If the cursor was

not at the appropriate depth and overlapped with the target, the target would remain light purple. Table 6 contains a list of the additional visual feedback elements.

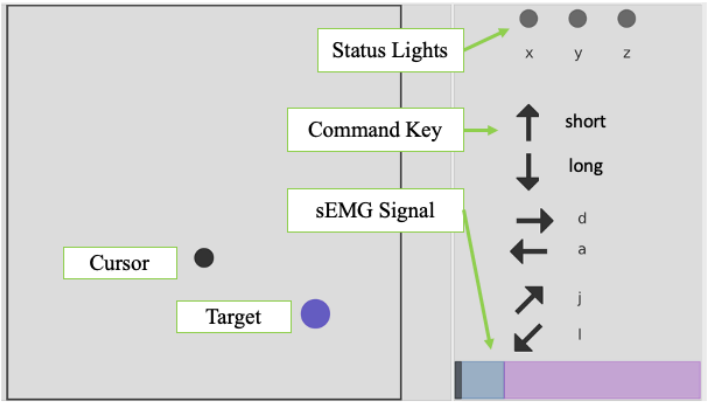



*Figure 23.* User Interface. The Task Interface is on the left-hand side, and the right-hand side contains the Information Interface.

The subjects completed 80 trials of the cursor-to-target task on the task interface. The task interface dimensions were normalized to have horizontal and vertical bounds of [-1,1] on a square, right-hand Cartesian coordinate system. At the beginning of each trial, the cursor started at the origin (i.e., center of the screen) with a diameter of 0.100. The cursor and target diameters were relative to the task interface dimensions. There were 16 unique target positions based on the target angle (45°, 135°, 225°, and 315°) and target diameter (0.025, 0.050, 0.150, 0.175), and all targets were positioned 0.80 units from the origin. A Block consisted of 4 trials and the subjects received a minimum 30 s break before starting the next Block; subjects could ask to rest longer. A set of four Blocks, or a Test, covered the 16 unique target positions in a pseudorandomized order. The five Tests each had the same pseudorandomized order of trials. The trials timed out after 30 s, which was based on preliminary testing. The total number of trials was determined by reviewing the results from *Chapters 2* and *3*, as well as our decision to reasonably minimize the duration of the in-person portion of the experiment due to coronavirus transmission concerns.

*Table 6.* Additional Visual Feedback During Task

| Object | Nominal Color | Feedback Color | Purpose |
|---|---|---|---|
| Cursor | Black | Dark gold | Indicates short input / up command |
| | | Bright blue | Indicates long input / down command |
| Target | Light purple | Green | Cursor at target's depth |
| | | Orange | Cursor selected target |

## 5.3 Analysis

### *5.3.1 Coordination Metrics*

To assess coordination between the hands and leg driving the robotic hand, we calculated a coordination score and DOF activation. The coordination score was calculated for each trial and was the weighted average of the time spent with 0-, 1-, 2-, and 3-DOFs activated. For each time step, points were added in proportion to the number of active DOFs. For example, if there were no DOFs active the time step would be assigned zero points. If any 1-DOF was active, then one point would be awarded. Each additional DOF earned another point. A final coordination score of 1 meant that on average 1-DOF was active during the trial. Larger coordination scores would indicate better and more coordinated performance. The coordination score did not detail which DOFs were active, only the number of DOFs. The DOF activation metric provided a finer analysis of which DOFs were used and the relative percentage for each trial. The DOF activation metric measured the proportion of the trial time spent with 0-, 1-, 2-, and/or 3-DOFs activated. This metric was calculated for the overall comparison between all DOFs, as well as within 1-DOF and 2-DOFs. For the DOF activation within 1-DOF, the metric assessed the proportion of the trial time that the *x*, *y*-, and *z*- axes were activated. The 2-DOF activation metric measured the proportion of trial time for *xy*, *xz*, and *yz* activation. These metrics were calculated for all trials, regardless of success.

### *5.3.2 Task Metrics*

Two task metrics were used to evaluate performance, percent of successful trials and completion time. The percent of successful trials was calculated as the number of successful trials divided by the total number of

trials in a Test (16 trials). The completion time was only calculated for successful trials. Both metrics were averaged over each Test.

## 5.4 Results

The coordination metrics and task metrics were calculated for all trials regardless of success, except for completion time. Performance during unsuccessful trials may provide interesting insights different than successful trials, however, the number of unsuccessful trials decreased substantially starting in Test 3. Across subjects, there were 25 unsuccessful trials (40% of the total trials) in Test 1 and 9 unsuccessful trials (14% of the total trials) in Test 2. The remaining Tests had a total unsuccessful trial count of 8 trials. Therefore, it was difficult to make meaningful comparisons between successful and unsuccessful trials.

### *5.4.1 Coordination Metrics*

The coordination score was calculated for each trial and averaged over all subjects per Test. As shown in Figure 24, the coordination score increased each subsequent Test with a slight decrease in the last Test. The average coordination scores ($\mu \pm \sigma$) in order of Test were $1.61 \pm 0.37$, $1.76 \pm 0.27$, $1.86 \pm 0.26$, $1.96 \pm 0.23$, and $1.89 \pm 0.25$. The coordination scores ranged from 0.61 in Test 1 to 2.42 in Test 5. The individual subject scores (averaged over all their trials) ranged from 1.65 to 1.93. The subjects did have a variety of experiences that may have been helpful for this task, such as prior myoelectric control and video game experience. Interestingly, the one subject with prior myoelectric control experience did not have the highest average coordination score. The subjects appeared to perform similar to each other in terms of the coordination score. There did not appear to be noticeable changes in the coordination score per Test when the results were disaggregated by trial success, target angle, or target depth. The slight decrease in the coordination score from Test 4 to Test 5 did not appear to be attributed to muscle fatigue. Two subjects reported feeling some muscle fatigue in their feet, but their maximum voluntary contractions decreased by less than 12% from before the task to after the task and remained well above the threshold.

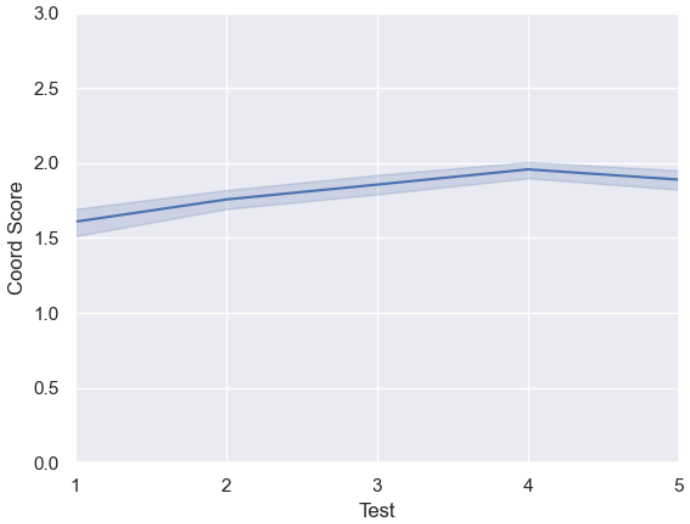



*Figure 24.* Coordination Score by Test. The coordination score is averaged across all subjects for each Test (16 trials). The shaded region shows the 95% confidence interval.

The coordination score improvements may be explained by increased 3-DOF activation and decreased 1-DOF activation, which was captured by the DOF activation metric (see Figure 25). The largest changes occurred for 1-DOF and 3-DOF activation between Test 1 and Test 4. The 1-DOF activation decreased by 14 percentage points and 3-DOF activation increased by 16 percentage points, which may be explained as the subjects decreasing their 1-DOF control and learning to increase their 3-DOF control. In contrast, the 0-DOF and 2-DOF activation remained relatively constant between Tests 1 and 4 with a difference in percentage points of 3 and 1, respectively. At most, these differences in percentages for 0-DOF and 2-DOF activation account for less than 1 s of the trial. The changes across DOF activations from Test 4 to Test 5 translate to a time difference of less than 0.5 s. Therefore, the most meaningful changes in DOF activations occurred from Test 1 to Test 4 for the 1-DOF and 3-DOF activations. As with the coordination score, there did not appear to be noticeable changes in the coordination score per Test when the results were disaggregated by trial success, target angle, or target depth.

In addition to the overall activation proportion between DOFs, it was also of interest to understand which axes were active within DOFs. As seen in Figure 25, about half of the trial time was spent with 2-DOFs activated. Additionally, 1-DOF was activated between 16% and 30% of the trial time. Together the

1-DOF and 2-DOF activations accounted for most of the trial time. Within the 1-DOF activation during a trial, the activation tended to originate from one of the hands with keyboard inputs (see Figure 26A). It was relatively uncommon for leg activation for 1-DOF activation, which occurred less than 10% of the time. For 2-DOF activation, the most likely pairing was for both hands together (*xz*-axes) for approximately 80% of the time (see Figure 26B). Coordination between the leg and left hand (*xy*-axes) occurred approximately 16% of time in 2-DOF activation, whereas the leg and right hand (*xz*-axes) coordination accounted for less than 4% of 2-DOF activation time. This analysis increased understanding of which axes were activated within a particular DOF activation case. However, it was also of interest to compare all combinations of DOFs and axes to determine an overall activation ranking. Each possible activation combination was ranked and percentage differences less than 1% were not considered different due to the small absolute time differences. The resulting 2-DOF activation order was both hands (*xz*, 41%), all three limbs (*xyz*, 19%), the left hand (*x*, 12%), and the left hand-leg (*xy*), right hand (*z*), and no activation tied and each accounted for 8%. The right hand-leg (*yz*) and leg only (*y*) were 2% of the time each. These results indicated that subjects tended to coordinate their hands followed by all three limbs.

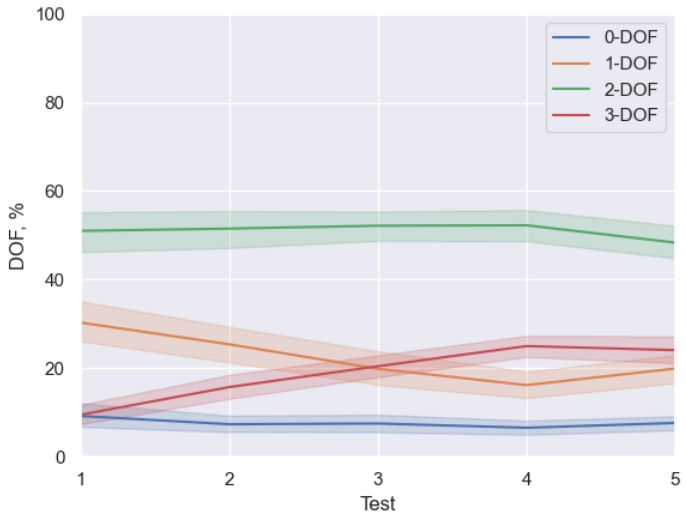



*Figure 25.* DOF Activation by Test. The DOF activation metric is averaged across each Test (16 trials). The shaded region shows the 95% confidence interval.

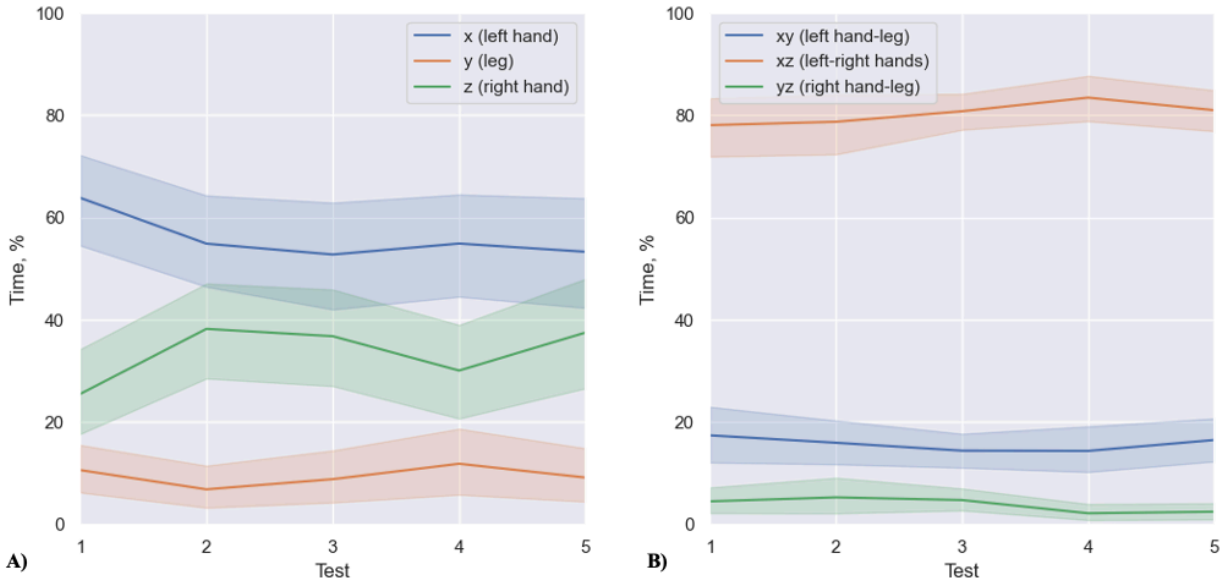



*Figure 26.* 1-DOF and 2-DOF Activation Breakdowns. The percentage of time is within either the 1-DOF (A) or 2-DOF (B) activation time and not for the entire trial time. The shaded area shows the 95% confidence interval. The results are averaged over all trials within a Test (16 trials).

### *5.4.2 Task Metrics*

The percent of successful trials was calculated for each Test and generally showed improvement throughout the experiment (see Figure 27). By Test 3, all subjects successfully completed more than 90% of the trials, and there were no unsuccessful trials in Test 4. The trend followed the coordination score, where there was a small decrease in the percent of successful trials for Test 5. There appeared to be relatively large differences in this metric between subjects for Tests 1 and 2. One subject only successfully completed one trial in Test 1 and seven trials in Test 2. Another subject successfully completed seven trials in Test 1 and then successfully completed all trials in the remaining Tests. The early differences in percent of successful trials may provide evidence of individual learning rates, however, the number of subjects was too low to correlate with demographic factors. These results indicated that the number of trials was sufficient for learning how to complete the task.

The completion time continually decreased on average in each subsequent Test (see Figure 28). The completion time was only calculated for successful trials, and trials had a maximum time of 30 s. In Test 1, the subjects used most of the allotted trial time with an average completion time of 21.98 ± 3.85 s

($\mu \pm \sigma$). The completion time decreased to less than half of the maximum trial time by Test 5 ($11.84 \pm 1.62$ s). Although completion time looked to be leveling off, it was not clear if the completion time plateaued in Test 5 and it could be interesting to observe this metric over more trials. These results combined with the percentage of successful trials supported the reduced trial times compared to *Chapter 3* (60 s maximum trial time).

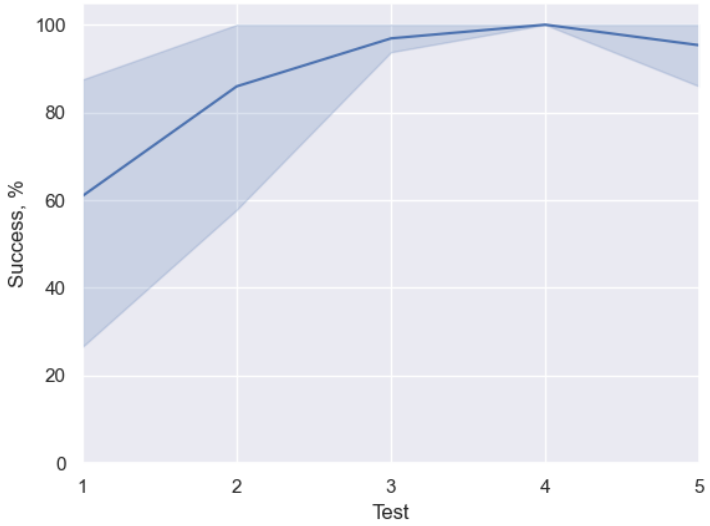



*Figure 27.* Percent of Successful Trials. Results are averaged over the 16 trials within a Test. The shaded region indicates the 95% confidence interval.

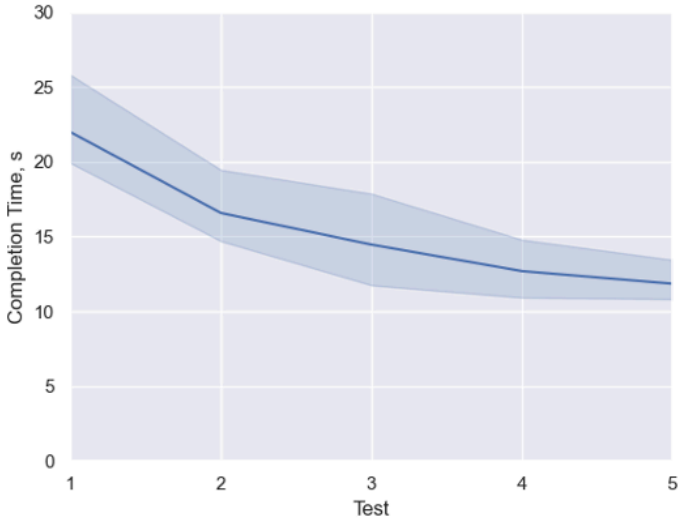



*Figure 28. Completion Time by Test.* The shaded area indicates the 95% confidence interval.

### *5.4.3 Additional Analysis*

The coordination metrics and task metrics provided a high-level assessment of the average subject performance; however, it was of interest to examine trial data in more detail. After reviewing data from low- and high-performance trials, one case was selected to highlight some of the changes in subject cursor control. The data shown in Figure 29 compares the normalized cursor velocities in each axis for the same target in Test 1 (Figure 29A) and Test 5 (Figure 29B). In the Test 1 trial, the subject moved the cursor primarily one axis at a time, except towards the end of the trial. The trial was unsuccessful and timed out with no inputs in the *z*-axis (i.e., depth/in and out motion). Directional corrections were made in both the *x*- and *y*- axes, as evidenced by the changing sign of the cursor velocities. In contrast, the trial data from Test 5 showed precisely timed inputs with minimal corrections. The subject correctly estimated the duration for the *y*-axis input, which was myoelectric control. The *x*-axis motion continued to the right once initiated with a gradual decrease and a minor correction to increase the velocity. The motion in the *z*-axis had the latest onset and smallest velocity but appeared to be precisely estimated to reach the target depth. The subject successfully completed the trial in 7.13 s. The subject may have been able to improve performance by holding the *x*-axis cursor velocity at the maximum for longer, ramping down the velocity quicker, and having an earlier onset for the *z*-axis cursor motion. Compared to the Test 1 trial data, it was evident that the subject improved in their precision and estimation of the cursor velocity in each axis. This subject had the highest coordination score averaged across all Tests and provided an example of the improvements that can occur throughout the experiment.

The subjects were provided with instructions that explicitly stated the goal of 3-DOF coordination. However, prior to the pilot study, it was not clear if subjects would aim to continually improve their coordination, or if they would choose to maximize the cursor velocity and move in 1-DOF at a time. The coordination score results (see Figure 24) and DOF activation results (see Figure 25) confirmed that subjects did try to improve their coordination. Concurrently, their percent of successful trials increased, and completion time decreased. We tested the correlation between completion time and coordination score using the Pearson correlation coefficient, $r$, for all successful trials. The resulting value, $r = -0.50$,

indicated a moderate, negative correlation between completion time and coordination score (see Figure 30). Increased coordination scores had a moderate tendency to also correlate with decreased completion times. These results provide evidence that increased coordination also correlated with increased efficiency when completing the task.

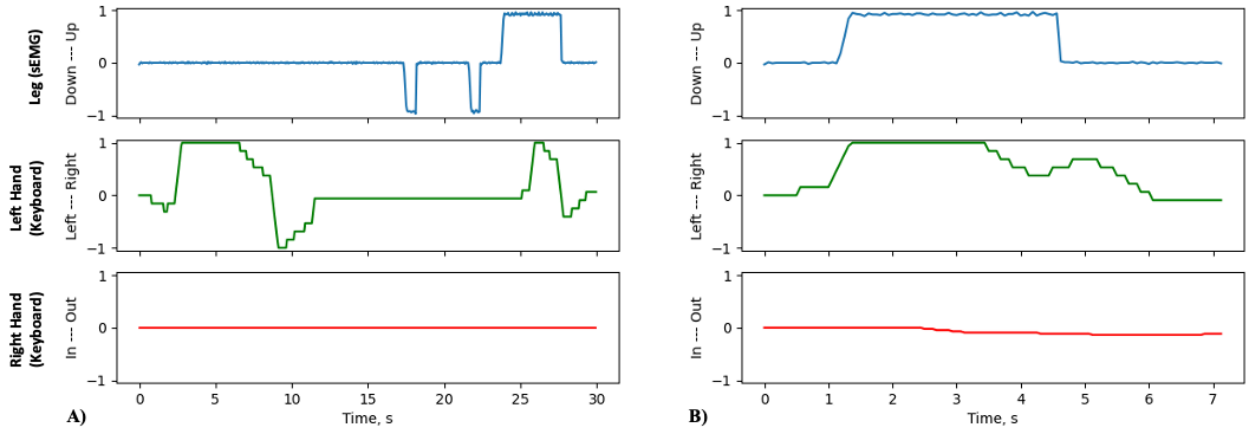



*Figure 29.* Cursor Velocities During a Trial. Normalized cursor velocities are shown for each axis for the selected trial (diameter = 0.05, angle = 45°) in Test 1 (A) and Test 5 (B) for the same subject. The trial in Test 1 was not successful.

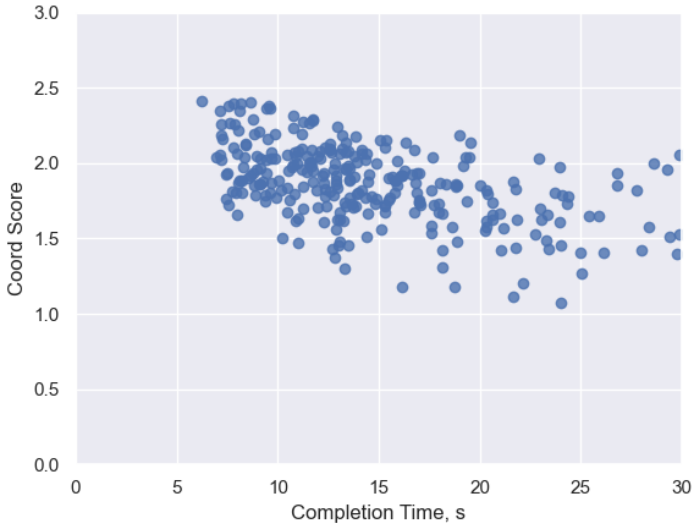



*Figure 30.* Completion Time versus Coordination Score for all Successful Trials.

## 5.5 Discussion

The results from this study showed increased coordination and task performance over its duration. The subjects were instructed to aim for 3-DOF coordination. The average coordination score in Test 4 and 5 was approximately a 2, which can be interpreted as an average of 2-DOF control during the trial. The goal was for subject to achieve 3-DOF coordination, and the highest coordination score was a 2.42 in Test 5. The average 3-DOF activation was 19% of the trial time. All subjects responded affirmatively when asked if they felt that "…you coordinated your hands and leg well." It may be beneficial in a future study to encourage subjects to increase their 3-DOF coordination. This encouragement could be achieved by showing the subjects their coordination score during or after each trial to give them a quantitative assessment of their performance. Another option would be to only allow cursor motion when all 3-DOFs are active as a training tool. This restriction could be removed during a subset of trials to evaluate their 3-DOF coordination retention.

The cursor and target both employed color feedback to provide additional information to the subjects during the task. Three of the subjects felt that the cursor color changes helped them confirm the command selection prior to cursor motion. The remaining subject said they largely ignored the color feedback and focused on resulting cursor motion. All subjects agreed that the target color feedback for depth aided them in aligning their cursor with the target. One subject further detailed that the target color feedback enabled them to time their actions. In contrast, the subjects minimally used the status lights and sEMG signal bar on the Information Interface. This information was not particularly surprising since it was expected that the subjects would allocate most of their visual attention to the Task Interface. However, the additional information was available as needed and the subjects did use the information occasionally to better understand myoelectric control (e.g., confirmation of crossing the threshold) and axes statuses.

This pilot study provided insights regarding the task design. The number of trials and trial duration was selected based on preliminary testing and the results from *Chapters 2* and *3*. The results from the task metrics generally support these selections for the given task design. As discussed, it would be of interest to

revise the training design to encourage increased 3-DOF coordination as compared to the modest improvements in coordination seen in this study. A task with direct, physical interaction may provide additional motivation and more obvious connection to the robot. Another change that should be considered in future studies is the muscle site. The tibialis anterior was originally selected due to its good signal quality and use in prior experiments (e.g., [102]). Subjects primarily flexed their foot or raised their toes to activate the muscle, and some reported fatigue in their foot. Overall, this study provided initial results and insights for myoelectric control of a collaborative robot during a computer-based task.

## 5.6 Conclusions

This pilot study expanded upon our previous work and demonstrated the sEMG system with a collaborative robot. The knowledge gained from previous studies helped us devise a task with a reasonable level of difficulty. Subjects appeared to follow the instructions and tried to coordinate their inputs, but the 3-DOF coordination remained relatively low. Different options have been discussed for encouraging increased 3-DOF coordination in future studies. Unlike previous chapters, the velocity controlled by sEMG remained constant. It would be technically possible to implement a proportional scheme, but some initial work should be done to better understand the relationship between the sEMG signal characteristics, the communication lag time, and the robot's acceleration settings. The robot performed in a reliable and safe manner, which built confidence for future experiments with direct interaction. This study focused on coordination and learning how to integrate a collaborative robot with myoelectric control. The metrics centered on coordination and task performance, and the other metrics we have previously used to address interaction factors like trust and cognitive workload should be incorporated in future studies. Although the robot did not physically interact with the subjects, the robot was integrated with the experiment software and established a basic infrastructure for future robotic myoelectric control experiments. Overall, this pilot study provided some interesting and encouraging results upon which to build more complex experiments in the future.

# Chapter 6

# Summary

The studies described in this dissertation explored the use of a single-site sEMG system at three different muscle sites for communication with a computer and robot that could be used for different applications. In *Chapter* 2, we developed two cursor control methods intended for high-level spinal cord injury patients, where available muscle sites were limited to the head and neck areas. For patient comfort, it was desirable to minimize the number of sensors and their intrusiveness while maximizing the functionality of the control method. Accordingly, both of our control methods used a single, noninvasive, surface, differential EMG sensor at the temporalis muscle that controlled a cursor in two DOFs. Fourteen inexperienced able-bodied subjects completed a Fitts's law-based cursor-to-target task using both control methods ("auto-rotate" and "manual rotate"). Subjects evaluated and compared performance between the two control methods, which both enabled the cursor to move in 2-DOFs. The control methods differed in the level of user control, where one allowed direction manipulation in 1-DOF and 2-DOFs in the other. Subjects also completed pre- and post-session surveys and the NASA-TLX for cognitive workload assessment. In general, subjects' performance improved with subsequent sessions within each control method. Subjects achieved a higher throughput (better performance) in the auto-rotate method, had lower workload scores, and tended to prefer this control method. However, about half the subjects felt the manual rotate method allowed them more control over cursor behavior. Our results suggested that a viable cursor control method can be achieved with only a single muscle site.

After demonstrating the feasibility of this sEMG control method, we increased the number of commands from two to four in *Chapter* 3. We developed a control task driven by subject-activated sEMG signals to observe the effects of automated training methodologies on performance, cognitive workload, and trust. Forty-eight subjects learned to use a sEMG-based command system to perform a Fitts's law-based cursor-to-target task with 120 training trials and 40 evaluation trials. Subjects were divided into four

groups: Control, Concurrent Feedback, Terminal Feedback, and Adaptive Threshold. The Control group trained and learned through repetition using the visual feedback of the cursor position. The Concurrent Feedback group received additional concurrent visual feedback during command input, and the Terminal Feedback group had supplementary visual feedback after command input. The Adaptive Threshold group did not have any additional feedback, but experienced changes in the cursor control designed to induce adaptation. Our results indicated that 1) additional concurrent and terminal visual feedback improved task performance, decreased cognitive workload, and increased trust during training, and 2) the groups converged in their command proficiency by the end of training.

*Chapter 3* established how to train novice subjects on this expanded sEMG system in a single session. However, there was an outstanding question regarding how to improve performance in trained subjects. The aim of the study in *Chapter 4* was to apply classifiers and adaptive features to sEMG signals obtained from subjects who were trained to perform cursor-control tasks (*Chapter 4* reused data collected in *Chapter 3*). We had developed a robust sEMG command system that relied on training the human rather than training a classifier, thereby enabling the human to correct for signal changes. The purpose of the study was to understand whether adding an adaptive timing feature and classifiers to our sEMG system could improve the performance of trained subjects. A Command Accuracy Test assessed subject proficiency at producing commands when prompted. The command classification accuracy was calculated for a control condition and two conditions that reflected possible adaptive features: the timing between sEMG signal inputs and a subject-specific classifier. The overall results showed significant improvements in command classification accuracy for both adaptive components compared to the control ($p < 0.0001$). However, some initially high performing subjects did not receive as much benefit. These results suggested that customizing the sEMG command system for individual subjects could improve their performance. Future work should investigate the effect of customizing the system for performance and co-adaptation, as well as using the adaptive features as a training tool to further improve command accuracy and efficiency.

For the final study (*Chapter 5*), we incorporated the knowledge gained in the previous studies to a three-handed coordination task with a collaborative robot. Four subjects participated in this pilot study and

completed 80 cursor-to-target trials. The subjects used their two natural hands to control a cursor on a screen in 2-DOFs and used their leg with sEMG to control the robotic hand. The robotic hand controlled the cursor in the third DOF. We calculated two metrics to assess coordination, the coordination score and the DOF activation. Subjects improved in the coordination score and DOF activation throughout the study duration. The subjects increased the percentage of trial time with 3-DOFs active and correspondingly decreased 1-DOF activation. The results indicated that subjects learned how to improve their coordination, while successfully completing trials and decreasing their trial completion time. The subjects tended to coordinate most of the time with their hands (41%) followed by all three hands (19%). Future studies should focus on increasing the proportion of 3-DOF coordination.

Overall, these studies showed that subjects were able to use the sEMG system by activating muscles at their temple, forearm, and lower leg. We demonstrated that the sEMG system can be expanded from two to four commands, and the efficacy of using concurrent feedback to train users on the expanded commands. The cursor-to-target tasks provided a consistent paradigm to assess the subjects' performances with various metrics related to efficiency and accuracy. These studied established that subjects can learn less intuitive myoelectric command systems in a single session, demonstrating that allocating relatively more learning to the human as opposed to the machine is a viable strategy. Future studies should explore how to optimize performance over multiple sessions and assess the abilities of the subjects to adapt to signal non-stationaries that occur session-to-session. Overall, this dissertation research has provided evidence of human learning and training for single-site sEMG, and the infrastructure for future research investigations.

REFERENCES


[1] S. M. O'Meara, M. C. Shyr, K. R. Lyons, and S. S. Joshi, "Comparing two different cursor control methods which use single-site surface electromyography," in *2019 9th international IEEE/EMBS conference on neural engineering (NER)*, 2019, pp. 1163–1166.

[2] S. M. O'Meara, J. Karasinski, C. L. Miller, S. Joshi, and S. K. Robinson, "The Effects of Training Methodology on Performance, Workload, and Trust During Human Learning of a Computer-Based Task," in *AIAA Scitech 2020 Forum*, 2020, p. 1110.

[3] S. M. O'Meara, J. A. Karasinski, C. L. Miller, S. S. Joshi, and S. K. Robinson, "Effects of Augmented Feedback and Motor Learning Adaptation on Human–Automation Interaction Factors," *Journal of Aerospace Information Systems*, pp. 1–14, 2021, doi: https://doi.org/10.2514/1.I010915.

[4] S. M. O'Meara, S. K. Robinson, and S. S. Joshi, "Classifiers and Adaptable Features Improve Myoelectric Command Accuracy in Trained Users," in *2021 10th International IEEE/EMBS Conference on Neural Engineering (NER)*, 2021, pp. 924–927. doi: https://doi.org/10.1109/NER49283.2021.9441256.

[5] M. Casadio, R. Ranganathan, and F. A. Mussa-Ivaldi, "The body-machine interface: a new perspective on an old theme," *Journal of Motor behavior*, vol. 44, no. 6, pp. 419–433, 2012.

[6] A. Duprès, R. N. Roy, S. Scannella, and F. Dehais, "Pre-stimulus EEG engagement ratio predicts inattentional deafness to auditory alarms in realistic flight simulator," in *3rd International Mobile Brain/Body Imaging Conference (MOBI 2018)*, 2018, pp. 1–3.

[7] F. Dehais *et al.*, "Monitoring Pilot's Cognitive Fatigue with Engagement Features in Simulated and Actual Flight Conditions Using an Hybrid fNIRS-EEG Passive BCI," in *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2018, pp. 544–549.

[8] T. Gateau, H. Ayaz, and F. Dehais, "In silico versus over the clouds: On-the-fly mental state estimation of aircraft pilots, using a functional near infrared spectroscopy based passive-BCI," *Frontiers in human neuroscience*, vol. 12, p. 187, 2018.

[9] G. Di Flumeri *et al.*, "Brain–computer interface-based adaptive automation to prevent out-of-the-loop phenomenon in air traffic controllers dealing with highly automated systems," *Frontiers in human neuroscience*, vol. 13, p. 296, 2019.

[10] B. Choi, D. Kim, M. Kim, and S. Hwang, "Effectiveness of EMG in Development of G-induced Loss of Consciousness (G-LOC) Warning System," in *Proceedings of the International Conference on Biomedical Engineering and Systems*, 2014, pp. 129-1–4.

[11] S. Kim, T. Cho, Y. Lee, H. Koo, B. Choi, and D. Kim, "G-LOC Warning Algorithms Based on EMG Features of the Gastrocnemius Muscle," *Aerospace medicine and human performance*, vol. 88, no. 8, pp. 737–742, Aug. 2017, doi: 10.3357/AMHP.4781.2017.

[12] M. F. Harrison, J. P. Neary, W. J. Albert, D. W. Veillette, N. P. McKenzie, and J. C. Croll, "Helicopter cockpit seat side and trapezius muscle metabolism with night vision goggles," *Aviation, space, and environmental medicine*, vol. 78, no. 10, pp. 995–998, Oct. 2007.

[13] C. G. de Oliveira and J. Nadal, "Back muscle EMG of helicopter pilots in flight: effects of fatigue, vibration, and posture," *Aviation, space, and environmental medicine*, vol. 75, no. 4, pp. 317–322, Apr. 2004.

[14] D. J. Hewson, P. J. McNair, and R. N. Marshall, "Aircraft control forces and EMG activity: comparison of novice and experienced pilots during simulated take-off and landing," *Aviation, space, and environmental medicine*, vol. 70, no. 8, pp. 745–751, Aug. 1999.

[15] C. Jorgensen, K. Wheeler, and S. Stepniewski, "Bioelectric Control of a 757 Class High Fidelity Aircraft Simulation," p. 10.

[16] M. Macchini, F. Schiano, and D. Floreano, "Personalized telerobotics by fast machine learning of body-machine interfaces," *IEEE Robotics and Automation Letters*, vol. 5, no. 1, pp. 179–186, 2019.

[17] M. Macchini, T. Havy, A. Weber, F. Schiano, and D. Floreano, "Hand-worn Haptic Interface for Drone Teleoperation," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 10212–10218.

[18] J. Miehlbradt *et al.*, "Data-driven body–machine interface for the accurate control of drones," *Proceedings of the National Academy of Sciences*, vol. 115, no. 31, pp. 7913–7918, 2018.

[19] A. Menshchikov, D. Lopatkin, E. Tsykunov, D. Tsetserukou, and A. Somov, "Realizing Body-Machine Interface for Quadrotor Control Through Kalman Filters and Recurrent Neural Network," in *2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, 2020, vol. 1, pp. 595–602.

[20] L. Summerer, "Thinking tomorrows' space–Research trends of the ESA advanced concepts team 2002–2012," *Acta Astronautica*, vol. 95, pp. 242–259, 2014.

[21] M. Broschart, C. de Negueruela, J. del R. Millán, and C. Menon, "Augmenting astronaut's capabilities through brain-machine interfaces," in *Proceedings of the 20th International Joint Conference on Artificial Intelligence, Workshop on Artificial Intelligence for Space Applications*, 2007, no. CONF.

[22] R. Poli, C. Cinel, A. Matran-Fernandez, F. Sepulveda, and A. Stoica, “Towards cooperative brain-computer interfaces for space navigation,” in *Proceedings of the 2013 international conference on Intelligent user interfaces*, 2013, pp. 149–160.

[23] J. del R. Millàn, P. W. Ferrez, and T. Seidl, “Validation of brain–machine interfaces during parabolic flight,” *International review of neurobiology*, vol. 86, pp. 189–197, 2009.

[24] E. B. Coffey, A.-M. Brouwer, E. S. Wilschut, and J. B. van Erp, “Brain–machine interfaces in space: using spontaneous rather than intentionally generated brain signals,” *Acta Astronautica*, vol. 67, no. 1–2, pp. 1–11, 2010.

[25] E. Scheme, A. Fougner, Ø. Stavdahl, A. D. Chan, and K. Englehart, “Examining the adverse effects of limb position on pattern recognition based myoelectric control,” in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, 2010, pp. 6337–6340.

[26] A. G. Barszap, I. M. Skavhaug, and S. S. Joshi, “Effects of muscle fatigue on the usability of a myoelectric human-computer interface,” *Hum Mov Sci*, vol. 49, pp. 225–38, Oct. 2016, doi: 10.1016/j.humov.2016.06.009.

[27] L. Hargrove, K. Englehart, and B. Hudgins, “A training strategy to reduce classification degradation due to electrode displacements in pattern recognition based myoelectric control,” *Biomedical signal processing and control*, vol. 3, no. 2, pp. 175–180, 2008.

[28] I. Kyranou, S. Vijayakumar, and M. S. Erden, “Causes of Performance Degradation in Non-invasive Electromyographic Pattern Recognition in Upper Limb Prostheses,” *Front Neurorobot*, vol. 12, p. 58, 2018, doi: 10.3389/fnbot.2018.00058.

[29] A. D. Roche, H. Rehbaum, D. Farina, and O. C. Aszmann, “Prosthetic myoelectric control strategies: a clinical perspective,” *Current Surgery Reports*, vol. 2, no. 3, p. 44, 2014.

[30] N. V. Iqbal and K. Subramaniam, “A review on upper-limb myoelectric prosthetic control,” *IETE Journal of Research*, vol. 64, no. 6, pp. 740–752, 2018.

[31] P. K. Artemiadis and K. J. Kyriakopoulos, “An EMG-based robot control scheme robust to time-varying EMG signal features,” *IEEE Trans Inf Technol Biomed*, vol. 14, no. 3, pp. 582–8, May 2010, doi: 10.1109/TITB.2010.2040832.

[32] I. Hussain, G. Salvietti, G. Spagnoletti, and D. Prattichizzo, “The Soft-SixthFinger: a Wearable EMG Controlled Robotic Extra-Finger for Grasp Compensation in Chronic Stroke Patients,” *IEEE Robotics and Automation Letters*, vol. 1, no. 2, pp. 1000–1006, Jul. 2016, doi: 10.1109/LRA.2016.2530793.

[33] I. Hussain, G. Spagnoletti, G. Salvietti, and D. Prattichizzo, “An EMG Interface for the Control of Motion and Compliance of a Supernumerary Robotic Finger,” *Front. Neurorobot.*, vol. 10, 2016, doi: 10.3389/fnbot.2016.00018.

[34] M. Ison and P. Artemiadis, “Proportional myoelectric control of robots: muscle synergy development drives performance enhancement, retainment, and generalization,” *IEEE Transactions on Robotics*, vol. 31, no. 2, pp. 259–268, 2015.

[35] J. Chen, M. Glover, C. Li, and C. Yang, “Development of a user experience enhanced teleoperation approach,” in *2016 International Conference on Advanced Robotics and Mechatronics (ICARM)*, Macau, China, Aug. 2016, pp. 171–177. doi: 10.1109/ICARM.2016.7606914.

[36] I. Vujaklija, D. Farina, and O. C. Aszmann, “New developments in prosthetic arm systems,” *Orthop Res Rev*, vol. 8, pp. 31–39, 2016, doi: 10.2147/ORR.S71468.

[37] M. Yoshikawa, M. Mikawa, and K. Tanaka, “Real-time hand motion estimation using EMG signals with support vector machines,” in *2006 SICE-ICASE International Joint Conference*, 2006, pp. 593–598.

[38] R. Hiroki and M. Iwase, “Hand and finger control of myo-prosthesis based on motion discriminator and voluntary control,” in *2017 11th Asian Control Conference (ASCC)*, 2017, pp. 1361–1366.

[39] W. S. Noble, “What is a support vector machine?,” *Nature biotechnology*, vol. 24, no. 12, pp. 1565–1567, 2006.

[40] A. D. Bellingegni *et al.*, “NLR, MLP, SVM, and LDA: a comparative analysis on EMG data from people with trans-radial amputation,” *Journal of neuroengineering and rehabilitation*, vol. 14, no. 1, p. 82, Aug. 2017, doi: ARTN 82 10.1186/s12984-017-0290-6.

[41] I.-M. Skavhaug, R. Bobell, B. Vernon, and S. S. Joshi, “Pilot study for a brain-muscle-computer interface using the extensor pollicis longus with preselected frequency bands,” in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2012, pp. 1727–1731.

[42] K. R. Lyons and S. S. Joshi, “Paralyzed subject controls telepresence mobile robot using novel sEMG brain-computer interface: Case study,” in *2013 IEEE 13th International Conference on Rehabilitation Robotics (ICORR)*, 2013, pp. 1–6.

[43] J. Weisz, A. G. Barszap, S. S. Joshi, and P. K. Allen, “Single muscle site semg interface for assistive grasping,” in *2014 IEEE/RSJ international conference on intelligent robots and systems*, 2014, pp. 2172–2178.

[44] C. Perez-Maldonado, A. S. Wexler, and S. S. Joshi, “Two-dimensional cursor-to-target control from single muscle site sEMG signals,” *IEEE Trans Neural Syst Rehabil Eng*, vol. 18, no. 2, pp. 203–9, Apr. 2010, doi: 10.1109/TNSRE.2009.2039394.

[45] I. M. Skavhaug, K. R. Lyons, A. Nemchuk, S. D. Muroff, and S. S. Joshi, “Learning to modulate the partial powers of a single sEMG power spectrum through a novel human-computer interface,” *Hum Movement Sci*, vol. 47, pp. 60–69, Jun. 2016, doi: 10.1016/j.humov.2015.12.003.

[46] J. Weisz, P. K. Allen, A. G. Barszap, and S. S. Joshi, “Assistive grasping with an augmented reality user interface,” *Int J Robot Res*, vol. 36, no. 5–7, pp. 543–562, Jun. 2017, doi: 10.1177/0278364917707024.

[47] K. Nas, L. Yazmalar, V. Sah, A. Aydin, and K. Ones, “Rehabilitation of spinal cord injuries,” *World J Orthop*, vol. 6, no. 1, pp. 8–16, Jan. 2015, doi: 10.5312/wjo.v6.i1.8.

[48] C. G. Pinheiro, E. L. Naves, P. Pino, E. Losson, A. O. Andrade, and G. Bourhis, “Alternative communication systems for people with severe motor disabilities: a survey,” *Biomed Eng Online*, vol. 10, no. 1, p. 31, Apr. 2011, doi: 10.1186/1475-925X-10-31.

[49] M. J. Cler and C. E. Stepp, “Discrete Versus Continuous Mapping of Facial Electromyography for Human-Machine Interface Control: Performance and Training Effects,” *IEEE Trans Neural Syst Rehabil Eng*, vol. 23, no. 4, pp. 572–80, Jul. 2015, doi: 10.1109/TNSRE.2015.2391054.

[50] C. N. Huang, C. H. Chen, and H. Y. Chung, “Application of facial electromyography in computer mouse access for people with disabilities,” *Disabil Rehabil*, vol. 28, no. 4, pp. 231–7, Feb. 2006, doi: 10.1080/09638280500158349.

[51] M. R. Williams and R. F. Kirsch, “Evaluation of Head Orientation and Neck Muscle EMG Signals as Command Inputs to a Human-Computer Interface for Individuals With High Tetraplegia,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 16, no. 5, pp. 485–496, Oct. 2008, doi: 10.1109/Tnsre.2008.2006216.

[52] C. Choi, Y. Na, B. Rim, Y. Kim, S. Kang, and J. Kim, “An SEMG computer interface using three myoelectric sites for proportional two-dimensional cursor motion control and clicking for individuals with spinal cord injuries,” *Med Eng Phys*, vol. 35, no. 6, pp. 777–783, Jun. 2013, doi: 10.1016/j.medengphy.2012.08.006.

[53] T. E. of E. Britannica, “Morse Code,” *Encyclopedia Britannica*. Sep. 18, 2020. Accessed: Jul. 22, 2021. [Online]. Available: https://www.britannica.com/topic/Morse-Code

[54] K. R. Lyons and S. S. Joshi, “Real-time evaluation of a myoelectric control method for high-level upper limb amputees based on homologous leg movements,” in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2016, pp. 6365–6368.

[55] R. W. Soukoreff and I. S. MacKenzie, “Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts’ law research in HCI,” *Int J Hum-Comput St*, vol. 61, no. 6, pp. 751–789, Dec. 2004, doi: 10.1016/j.ijhcs.2004.09.001.

[56] S. G. Hart, “NASA task load index (TLX). Volume 1.0; Paper and pencil package,” 1986.

[57] M. R. Williams and R. F. Kirsch, “Case study: Head orientation and neck electromyography for cursor control in persons with high cervical tetraplegia,” *Journal of Rehabilitation Research & Development*, vol. 53, no. 4, 2016.

[58] C. Choi, S. Micera, J. Carpaneto, and J. Kim, “Development and quantitative performance evaluation of a noninvasive EMG computer interface,” *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 1, pp. 188–191, 2008.

[59] O. Safir, A. Dubrowski, Y. Hui, D. Backstein, and H. Carnahan, “Self-directed practice scheduling is equivalent to instructor guided practice when learning a complex surgical skill,” *Procedia-Social and Behavioral Sciences*, vol. 2, no. 2, pp. 792–796, 2010.

[60] J. Karasinski, S. Robinson, P. Handley, and K. Duda, “Real-time performance feedback in a manually-controlled spacecraft inspection task,” in *AIAA Modeling and Simulation Technologies Conference*, 2017, p. 1314.

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

[62] H. M. Bosse *et al.*, “The benefit of repetitive skills training and frequency of expert feedback in the early acquisition of procedural skills,” *BMC medical education*, vol. 15, no. 1, pp. 1–10, 2015.

[63] L. M. Al-Saud *et al.*, “Feedback and motor skill acquisition using a haptic dental simulator,” *European Journal of Dental Education*, vol. 21, no. 4, pp. 240–247, 2017.

[64] D. Cecilio-Fernandes, F. Cnossen, J. Coster, A. (Debbie) C. Jaarsma, and R. A. Tio, “The Effects of Expert and Augmented Feedback on Learning a Complex Medical Skill,” *Perceptual and motor skills*, vol. 127, no. 4, pp. 766–784, 2020.

[65] M. Huet, D. M. Jacobs, C. Camachon, O. Missenard, R. Gray, and G. Montagne, “The education of attention as explanation of variability of practice effects: learning the final approach phase in a flight simulator.,” *Journal of Experimental Psychology: Human Perception and Performance*, vol. 37, no. 6, p. 1841, 2011.

[66] Y. B. Eisma, C. Borst, R. van Paassen, and J. de Winter, “Augmented visual feedback: cure or distraction?,” *Human factors*, p. 0018720820924602, 2020.

[67] E. Triantafyllidis, C. Mcgreavy, J. Gu, and Z. Li, "Study of Multimodal Interfaces and the Improvements on Teleoperation," *IEEE Access*, vol. 8, pp. 78213–78227, 2020.

[68] F. Marschall, A. Bund, and J. Wiemeyer, "Does frequent augmented feedback really degrade learning? A meta-analysis," *Bewegung und Training*, vol. 1, pp. 75–86, 2007.

[69] D. G. Liebermann, L. Katz, M. D. Hughes, R. M. Bartlett, J. McClements, and I. M. Franks, "Advances in the application of information technology to sport performance," *Journal of sports sciences*, vol. 20, no. 10, pp. 755–769, 2002.

[70] D. R. Billings, "Efficacy of adaptive feedback strategies in simulation-based training," *Military Psychology*, vol. 24, no. 2, pp. 114–133, 2012.

[71] J. Hattie and H. Timperley, "The power of feedback," *Review of educational research*, vol. 77, no. 1, pp. 81–112, 2007.

[72] S. Wijewickrema, X. Ma, J. Bailey, G. Kennedy, and S. O'Leary, "Feedback Techniques in Computer-Based Simulation Training: A Survey," *arXiv preprint arXiv:1705.04683*, 2017.

[73] S. Dosen *et al.*, "Multichannel electrotactile feedback with spatial and mixed coding for closed-loop control of grasping force in hand prostheses," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 3, pp. 183–195, Mar. 2016, doi: 10.1109/TNSRE.2016.2550864.

[74] M. Markovic, H. Karnal, B. Graimann, D. Farina, and S. Dosen, "GLIMPSE: Google Glass interface for sensory feedback in myoelectric hand prostheses," *Journal of neural engineering*, vol. 14, no. 3, p. 036007, 2017.

[75] M. A. Schweisfurth, M. Markovic, S. Dosen, F. Teich, B. Graimann, and D. Farina, "Electrotactile EMG feedback improves the control of prosthesis grasping force," *Journal of neural engineering*, vol. 13, no. 5, p. 056010, 2016.

[76] A. W. Shehata, L. F. Engels, M. Controzzi, C. Cipriani, E. J. Scheme, and J. W. Sensinger, "Improving internal model strength and performance of prosthetic hands using augmented feedback," *Journal of neuroengineering and rehabilitation*, vol. 15, no. 1, p. 70, Jul. 2018, doi: ARTN 70 10.1186/s12984-018-0417-4.

[77] H. Huang, S. L. Wolf, and J. He, "Recent developments in biofeedback for neuromotor rehabilitation," *Journal of neuroengineering and rehabilitation*, vol. 3, no. 1, p. 11, Jun. 2006, doi: 10.1186/1743-0003-3-11.

[78] J. V. Basmajian, "Control and training of individual motor units," *Science*, vol. 141, no. 3579, pp. 440–441, Aug. 1963.

[79] R. Stanton, L. Ada, C. M. Dean, and E. Preston, “Biofeedback improves performance in lower limb activities more than usual therapy in people following stroke: a systematic review,” *J Physiother*, vol. 63, no. 1, pp. 11–16, Jan. 2017, doi: 10.1016/j.jphys.2016.11.006.

[80] A. R. Pritchett, “Aviation Automation: General Perspectives and Specific Guidance for the Design of Modes and Alerts,” *Reviews of Human Factors and Ergonomics*, vol. 5, no. 1, pp. 82–113, Jun. 2009, doi: 10.1518/155723409X448026.

[81] R. Y. T. Seow, S. Betts, and J. R. Anderson, “Transfer effects of varied practice and adaptation to changes in complex skill acquisition,” in *Proceedings of the 17th International Conference on Cognitive Modelling*, 2019, pp. 222–227.

[82] D. A. Braun, A. Aertsen, D. M. Wolpert, and C. Mehring, “Motor task variation induces structural learning,” *Current Biology*, vol. 19, no. 4, pp. 352–357, 2009.

[83] R. E. Johnson, K. P. Kording, L. J. Hargrove, and J. W. Sensinger, “Does EMG control lead to distinct motor adaptation?,” *Frontiers in neuroscience*, vol. 8, p. 302, 2014.

[84] R. K. Lyons and S. S. Joshi, “Effects of Mapping Uncertainty on Visuomotor Adaptation to Trial-By-Trial Perturbations with Proportional Myoelectric Control,” in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018, pp. 5178–5181.

[85] A. W. Shehata, E. J. Scheme, and J. W. Sensinger, “The effect of myoelectric prosthesis control strategies and feedback level on adaptation rate for a target acquisition task,” in *2017 International Conference on Rehabilitation Robotics (ICORR)*, 2017, pp. 200–204.

[86] D. Sternad, “It’s not (only) the mean that matters: variability, noise and exploration in skill learning,” *Current opinion in behavioral sciences*, vol. 20, pp. 183–195, 2018.

[87] S. G. Hart, “NASA-task load index (NASA-TLX); 20 years later,” 2006, vol. 50, pp. 904–908.

[88] S. G. Hart and L. E. Staveland, “Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research,” in *Advances in psychology*, vol. 52, Elsevier, 1988, pp. 139–183.

[89] A. H. Roscoe and G. A. Ellis, “A subjective rating scale for assessing pilot workload in flight: A decade of practical use,” Royal Aerospace Establishment Farnborough (United Kingdom), 1990.

[90] P. Jorna, “Heart rate and workload variations in actual and simulated flight,” *Ergonomics*, vol. 36, no. 9, pp. 1043–1054, 1993.

[91] S. M. Casner and B. F. Gore, “Measuring and evaluating workload: A primer,” *NASA Technical Memorandum*, vol. 216395, p. 2010, 2010.

[92] V. J. Gawron, *Human Performance, Workload, and Situational Awareness Measures Handbook, -2-Volume Set*. CRC Press, 2019.

[93] R. C. Williges and W. W. Wierwille, “Behavioral measures of aircrew mental workload,” *Human Factors*, vol. 21, no. 5, pp. 549–574, 1979.

[94] J. D. Lee and K. A. See, “Trust in Automation: Designing for Appropriate Reliance,” *Hum Factors*, vol. 46, no. 1, pp. 50–80, Mar. 2004, doi: 10.1518/hfes.46.1.50_30392.

[95] C. D. Wickens, K. Gempler, and M. E. Morphew, “Workload and reliability of predictor displays in aircraft traffic avoidance,” *Transportation Human Factors*, vol. 2, no. 2, pp. 99–126, 2000.

[96] K. A. Hoff and M. Bashir, “Trust in automation: Integrating empirical evidence on factors that influence trust,” *Hum Factors*, vol. 57, no. 3, pp. 407–434, May 2015, doi: 10.1177/0018720814547570.

[97] P. M. Fitts, “The information capacity of the human motor system in controlling the amplitude of movement,” *Journal of experimental psychology*, vol. 47, no. 6, p. 381, 1954.

[98] A. W. Salmoni, R. A. Schmidt, and C. B. Walter, “Knowledge of results and motor learning: a review and critical reappraisal.,” *Psychological bulletin*, vol. 95, no. 3, p. 355, 1984.

[99] K. Lyons and B. Margolis, “AxoPy: A Python Library for Implementing Human-Computer Interface Experiments,” *The Journal of Open Source Software*, vol. 4, p. 1191, 2019.

[100] J. A. Karasinski, S. K. Robinson, K. R. Duda, and Z. Prasov, “Development of real-time performance metrics for manually-guided spacecraft operations,” in *2016 IEEE Aerospace Conference*, Mar. 2016, pp. 1–9. doi: 10.1109/AERO.2016.7500734.

[101] K. D. Lyons *et al.*, “Autonomous Psychological Support for Isolation and Confinement,” *Aerospace Medicine and Human Performance*, vol. 91, no. 11, pp. 876–885, 2020.

[102] K. R. Lyons, S. S. Joshi, S. S. Joshi, and K. R. Lyons, “Upper Limb Prosthesis Control for High-Level Amputees via Myoelectric Recognition of Leg Gestures,” *IEEE Trans Neural Syst Rehabil Eng*, vol. 26, no. 5, pp. 1056–1066, May 2018, doi: 10.1109/TNSRE.2018.2807360.

[103] M. Chhabra and R. A. Jacobs, “Near-optimal human adaptive control across different noise environments,” *Journal of Neuroscience*, vol. 26, no. 42, pp. 10883–10887, Oct. 2006, doi: 10.1523/JNEUROSCI.2238-06.2006.

[104] J.-Y. Jian, A. M. Bisantz, and C. G. Drury, “Foundations for an empirically determined scale of trust in automated systems,” *International Journal of Cognitive Ergonomics*, vol. 4, no. 1, pp. 53–71, 2000.

[105] D. Bates, M. Mächler, B. Bolker, and S. Walker, “Fitting linear mixed-effects models using lme4,” *arXiv preprint arXiv:1406.5823*, 2014.

[106] M. Ison and P. Artemiadis, “Proportional myoelectric control of robots: muscle synergy development drives performance enhancement, retainment, and generalization,” *IEEE Transactions on Robotics*, vol. 31, no. 2, pp. 259–268, 2015.

[107] M. T. Dzindolet, S. A. Peterson, R. A. Pomranky, L. G. Pierce, and H. P. Beck, “The role of trust in automation reliance,” *Int J Hum-Comput St*, vol. 58, no. 6, pp. 697–718, 2003.

[108] H. Zhang, Y. Zhao, F. Yao, L. Xu, P. Shang, and G. Li, “An adaptation strategy of using LDA classifier for EMG pattern recognition,” in *2013 35th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, 2013, pp. 4267–4270.

[109] J. M. Hahne, S. Dähne, H.-J. Hwang, K.-R. Müller, and L. C. Parra, “Concurrent Adaptation of Human and Machine Improves Simultaneous and Proportional Myoelectric Control,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, no. 4, pp. 618–627, Jul. 2015, doi: 10.1109/TNSRE.2015.2401134.

[110] D. Yeung, D. Farina, and I. Vujaklija, “Directional Forgetting for Stable Co-Adaptation in Myoelectric Control,” *Sensors*, vol. 19, no. 9, p. 2203, 2019.

[111] S. I. Benchabane, N. Saadia, and A. Ramdane-Cherif, “Novel algorithm for conventional myocontrol of upper limbs prosthetics,” *Biomedical Signal Processing and Control*, vol. 57, p. 101791, 2020.

[112] F. Pedregosa *et al.*, “Scikit-learn: Machine learning in Python,” *the Journal of machine Learning research*, vol. 12, pp. 2825–2830, 2011.

[113] C. Beleites, U. Neugebauer, T. Bocklitz, C. Krafft, and J. Popp, “Sample size planning for classification models,” *Analytica chimica acta*, vol. 760, pp. 25–33, 2013.

[114] A. Ajoudani, A. M. Zanchettin, S. Ivaldi, A. Albu-Schaffer, K. Kosuge, and O. Khatib, “Progress and prospects of the human-robot collaboration,” *Auton Robot*, vol. 42, no. 5, pp. 957–975, Jun. 2018, doi: 10.1007/s10514-017-9677-2.

[115] D. A. Kurek and H. H. Asada, “The MantisBot: Design and impedance control of supernumerary robotic limbs for near-ground work,” in *Robotics and Automation (ICRA), 2017 IEEE International Conference on*, 2017, pp. 5942–5947.

[116] B. L. Bonilla and H. H. Asada, “A robot on the shoulder: Coordinated human-wearable robot control using coloured petri nets and partial least squares predictions,” in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, 2014, pp. 119–125.

[117] C.-Y. Shin, J. Bae, and D. Hong, “Ceiling work scenario based hardware design and control algorithm of supernumerary robotic limbs,” in *2015 15th International Conference on Control, Automation and Systems (ICCAS)*, 2015, pp. 1228–1230.

[118] V. Vatsal and G. Hoffman, “A Wearable Robotic Forearm for Human-Robot Collaboration,” in *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 2018, pp. 329–330.

[119] M. Y. Saraiji, T. Sasaki, K. Kunze, K. Minamizawa, and M. Inami, “MetaArms: Body Remapping Using Feet-Controlled Artificial Arms,” in *The 31st Annual ACM Symposium on User Interface Software and Technology - UIST ’18*, Berlin, Germany, 2018, pp. 65–74. doi: 10.1145/3242587.3242665.

[120] B. Llorens-Bonilla, F. Parietti, and H. H. Asada, “Demonstration-based control of supernumerary robotic limbs.,” 2012, pp. 7–12.

[121] F. Parietti and H. Asada, “Supernumerary Robotic Limbs for Human Body Support,” *IEEE Trans. Robot.*, vol. 32, no. 2, pp. 301–311, Apr. 2016, doi: 10.1109/Tro.2016.2520486.

[122] F. Parietti, K. Chan, and H. H. Asada, “Bracing the human body with supernumerary robotic limbs for physical assistance and load reduction,” 2014, pp. 141–148.

[123] F. Y. Wu and H. H. Asada, “Implicit and Intuitive Grasp Posture Control for Wearable Robotic Fingers: A Data-Driven Method Using Partial Least Squares,” *Ieee T Robot*, vol. 32, no. 1, pp. 176–186, Feb. 2016, doi: 10.1109/Tro.2015.2506731.

[124] I. Hussain, L. Meli, C. Pacchierotti, and D. Prattichizzo, “A soft robotic supernumerary finger and a wearable cutaneous finger interface to compensate the missing grasping capabilities in chronic stroke patients,” in *2017 IEEE World Haptics Conference (WHC)*, Jun. 2017, pp. 183–188. doi: 10.1109/WHC.2017.7989898.

[125] L. Tiziani, A. Hart, T. Cahoon, F. Y. Wu, H. H. Asada, and F. L. Hammond, “Empirical characterization of modular variable stiffness inflatable structures for supernumerary grasp-assist devices,” *Int J Robot Res*, vol. 36, no. 13–14, pp. 1391–1413, Dec. 2017, doi: 10.1177/0278364917714062.

[126] M. Ariyanto, R. Ismail, J. D. Setiawan, and Z. Arifin, “Development of low cost supernumerary robotic fingers as an assistive device,” in *2017 4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, 2017, pp. 1–6.

[127] S. Leigh and P. Maes, “Body integrated programmable joints interface,” in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2016, pp. 6053–6057.

[128] V. Vatsal and G. Hoffman, “Wearing your arm on your sleeve: Studying usage contexts for a wearable robotic forearm,” in *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 2017, pp. 974–980.

[129] S. W. Leigh, H. Agrawal, and P. Maes, “Robotic Symbionts Interweaving Human and Machine Actions,” *IEEE Pervas Comput*, vol. 17, no. 2, pp. 34–43, Jun. 2018, doi: Doi 10.1109/Mprv.2018.022511241.

[130] E. Abdi, M. Bouri, E. Burdet, S. Himidan, and H. Bleuler, “Positioning the endoscope in laparoscopic surgery by foot: Influential factors on surgeons’ performance in virtual trainer,” in *Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE*, 2017, pp. 3944–3948.

[131] E. Abdi, E. Burdet, M. Bouri, S. Himidan, and H. Bleuler, “In a demanding task, three-handed manipulation is preferred to two-handed manipulation,” *Sci Rep*, vol. 6, p. 21758, Feb. 2016, doi: 10.1038/srep21758.

[132] E. Abdi, E. Burdet, M. Bouri, and H. Bleuler, “Control of a Supernumerary Robotic Hand by Foot: An Experimental Study in Virtual Reality,” *Plos One*, vol. 10, no. 7, p. e0134501, 2015, doi: 10.1371/journal.pone.0134501.

[133] E. Velloso, D. Schmidt, J. Alexander, H. Gellersen, and A. Bulling, “The Feet in Human-Computer Interaction: A Survey of Foot-Based Interaction,” *Acm Comput Surv*, vol. 48, no. 2, p. 21, Nov. 2015, doi: Artn 21 10.1145/2816455.

[134] CDC, “Cases, Data, and Surveillance,” *Centers for Disease Control and Prevention*, Feb. 11, 2020. https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-discovery/hospitalization-death-by-age.html (accessed Jun. 18, 2021).

[135] “DocuSign | #1 in Electronic Signature and Agreement Cloud.” https://www.docusign.com/ (accessed Jun. 18, 2021).

[136] Office of the Commissioner, “Collection of Race and Ethnicity Data in Clinical Trials,” *U.S. Food and Drug Administration*, Nov. 08, 2019. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/collection-race-and-ethnicity-data-clinical-trials (accessed Jun. 18, 2021).

[137] L. J. Elias, M. P. Bryden, and M. B. Bulman-Fleming, “Footedness is a better predictor than is handedness of emotional lateralization,” *Neuropsychologia*, vol. 36, no. 1, pp. 37–43, Jan. 1998, doi: 10.1016/S0028-3932(97)00107-3.

[138] “UR5 collaborative robot arm | Flexible and lightweight cobot.” https://www.universal-robots.com/products/ur5-robot/ (accessed Jun. 18, 2021).

[139] “Real-Time Data Exchange (RTDE) Guide - 22229.” https://www.universal-robots.com/articles/ur/interface-communication/real-time-data-exchange-rtde-guide/ (accessed Jun. 18, 2021).

[140] *SDU Robotics / ur_rtde*. Accessed: Jun. 18, 2021. [Online]. Available: https://gitlab.com/sdurobotics/ur_rtde

[141] Photo by Jordan Whitt on Unsplash.

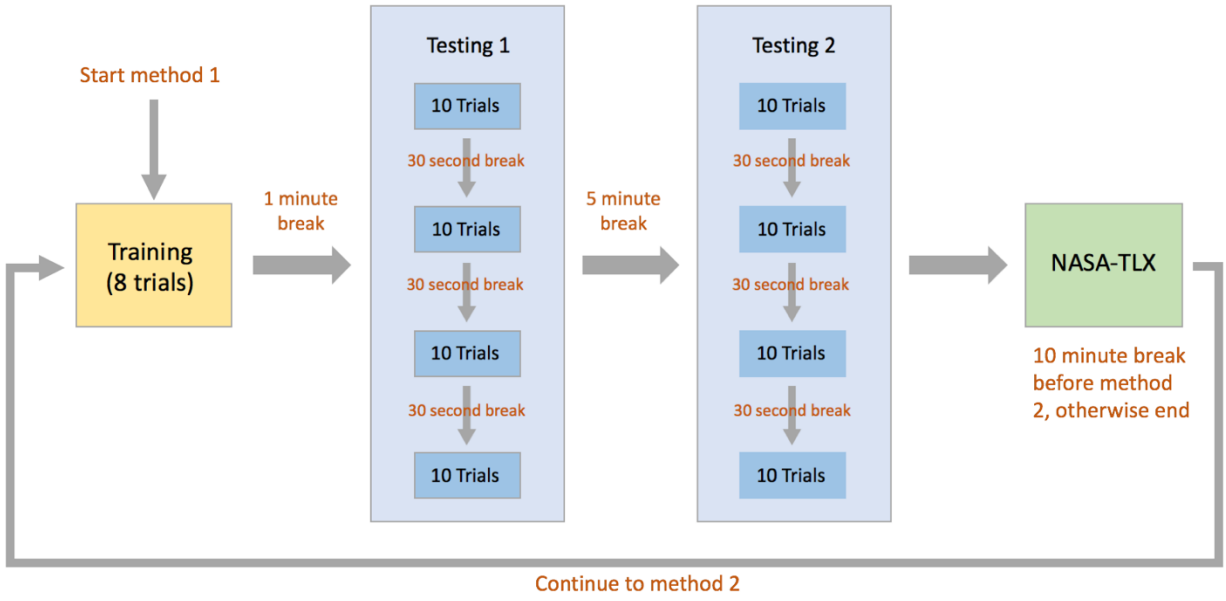
[142] “Bedford Scale.” http://web.mit.edu/16.400/www/flight_sim/BedfordScale.pdf (accessed Jul. 27, 2021).

APPENDICES

APPENDIX A

Experiment Protocol for Chapter 1

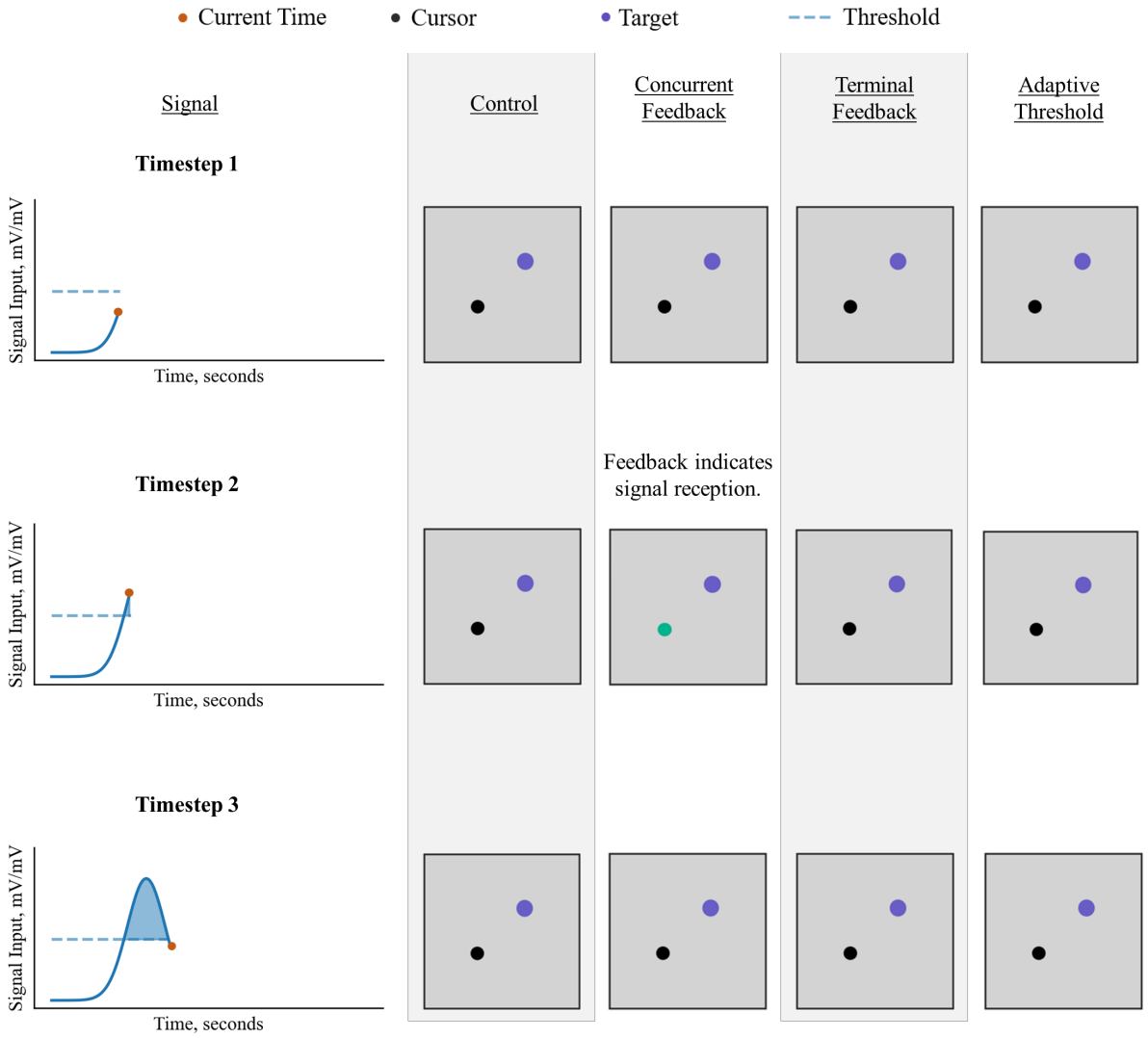
The diagram illustrates the experiment protocol for both the auto-rotate and manual rotate methods. Group 1 started with the auto-rotate method and then proceeded to use the manual rotate method. Group 2 used the manual rotate method before the auto-rotate method.

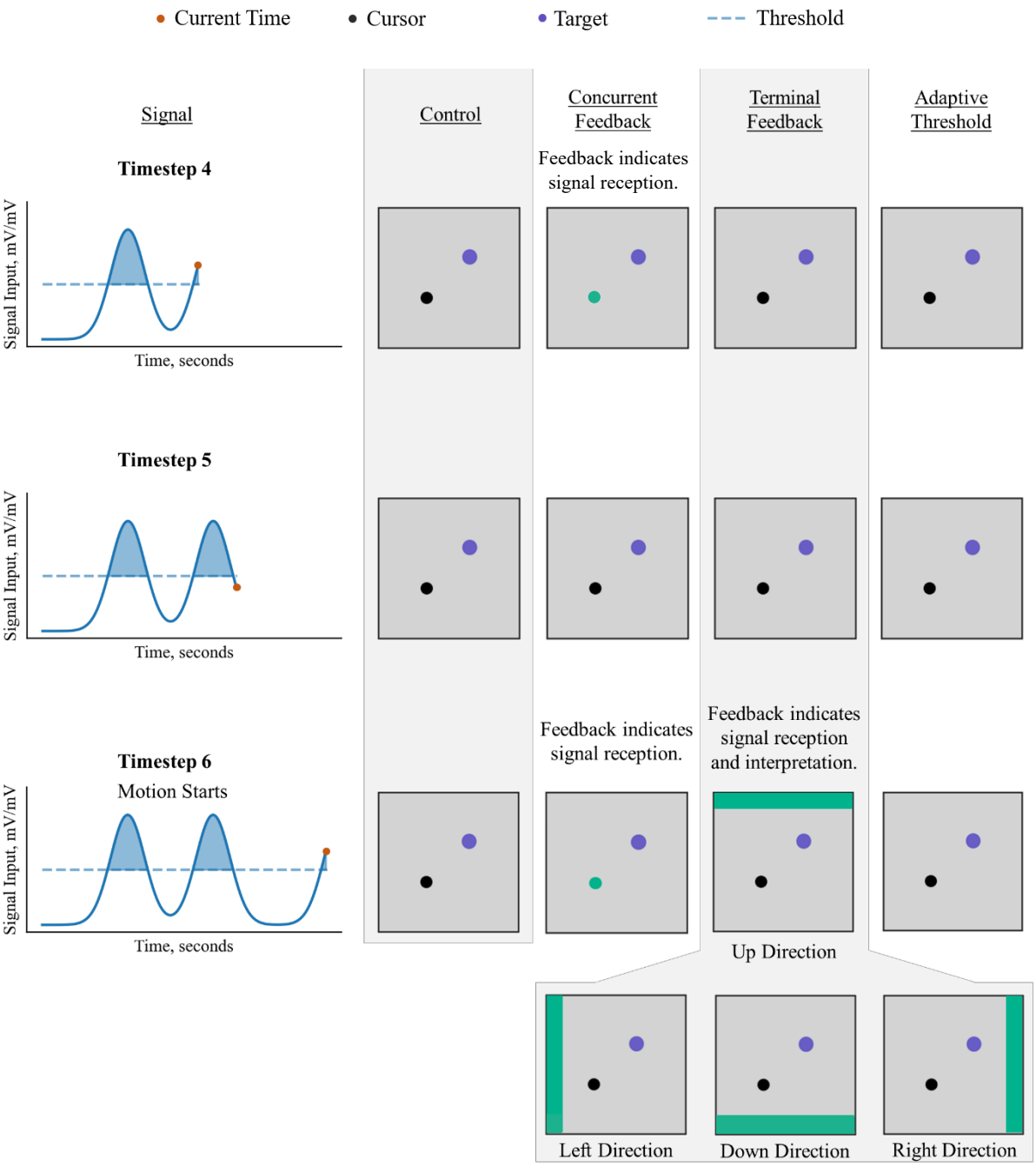


APPENDIX B

User Interface Displays for Different Training Methodologies for Chapter 3

The signal and user interface illustrates the relationship between the signal input at particular times and the resulting changes in the user interface. All groups are shown for completeness, although there are no dynamic, visual elements for the Control and Adaptive Threshold groups. The varying threshold for the Adaptive Threshold group is not shown in the signal plots. All possible visual elements are shown for the Terminal Feedback group. For better visibility of the cursor and target, the user interfaces are not to scale. The signal plot is generic and is not specifically showing short or long inputs.




Current Time
Cursor
Target
Threshold
Signal
Control
Concurrent Feedback
Terminal Feedback
Adaptive Threshold
Timestep 4
Signal Input, mV/mV
Time, seconds
Feedback indicates signal reception.
Timestep 5
Signal Input, mV/mV
Time, seconds
Timestep 6
Motion Starts
Signal Input, mV/mV
Time, seconds
Feedback indicates signal reception.
Feedback indicates signal reception and interpretation.
Up Direction
Left Direction
Down Direction
Right Direction

APPENDIX C

Modified Bedford Workload Scale for Chapter 3

The subjects refer to the below flow diagram [142] when rating their cognitive workload every 10 trials (a.k.a. a Block).