Delivering feedback during pre-device training facilitates the retention of novel myoelectric skills: a laboratory and home-based study

Citation for published version:

Digital Object Identifier (DOI):
10.1088/1741-2552/acc4ea

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Journal of Neural Engineering

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Delaying feedback during pre-device training facilitates the retention of novel myoelectric skills: a laboratory and home-based study

To cite this article before publication: Simon Ainslie Stuttaford et al 2023 J. Neural Eng. in press https://doi.org/10.1088/1741-2552/acc4ea

This Accepted Manuscript is © 2023 The Author(s). Published by IOP Publishing Ltd.

As the Version of Record of this article is going to be / has been published on a gold open access basis under a CC BY 4.0 licence, this Accepted Manuscript is available for reuse under a CC BY 4.0 licence immediately.

Everyone is permitted to use all or part of the original content in this article, provided that they adhere to all the terms of the licence https://creativecommons.org/licenses/by/4.0

Although reasonable endeavours have been taken to obtain all necessary permissions from third parties to include their copyrighted content within this article, their full citation and copyright line may not be present in this Accepted Manuscript version. Before using any content from this article, please refer to the Version of Record on IOPscience once published for full citation and copyright details, as permissions may be required. All third party content is fully copyright protected and is not published on a gold open access basis under a CC BY licence, unless that is specifically stated in the figure caption in the Version of Record.

View the article online for updates and enhancements.
Delaying feedback during pre-device training facilitates the retention of novel myoelectric skills: a laboratory and home-based study

Simon A. Stuttaford1  Sigrid S.G. Dupan2
s.stuttaford1@newcastle.ac.uk  sigrid.dupan@edinburgh.ac.uk

Kianoush Nazarpour2  Matthew Dyson1
kianoush.nazarpour@edinburgh.ac.uk  matthew.dyson@newcastle.ac.uk

1Intelligent Sensing Laboratory, School of Engineering, Newcastle University
2Edinburgh Neuroprosthetics Laboratory, School of Informatics, The University of Edinburgh

Abstract

Objective. The objective of this study was to assess the impact of delayed feedback training on the retention of novel myoelectric skills, and to demonstrate the use of this training approach in the home environment.

Approach. We trained limb-intact participants to use a motor learning-based upper-limb prosthesis control scheme called abstract decoding. A delayed feedback paradigm intended to prevent within-trial adaptation and to facilitate motor learning was used. We conducted two multi-day experiments. Experiment 1 was a laboratory-based study consisting of two groups trained over a four-day period with concurrent or delayed feedback. An additional follow-up session took place after 18 days to assess the retention of motor skills. Experiment 2 was a home-based pilot study that took place over five consecutive days to investigate delayed feedback performance when using bespoke training structures.

Main Results. Approximately 35,000 trials were collected across both experiments. Experiment 1 found that the retention of motor skills for the delayed feedback group was significantly better than that of their concurrent feedback counterparts. In addition, the delayed feedback group improved their retention of motor skills across days, whereas the concurrent feedback group did not. Experiment 2 demonstrated that by using a bespoke training protocol in an environment that is more conducive to learning, it is possible for participants to become highly accurate in the absence of feedback.

Significance. These results show that with delayed feedback training, it is possible to retain novel myoelectric skills. Using abstract decoding participants can activate four distinct muscle patterns without using complex algorithms. The accuracy achieved in the pilot study supports the feasibility of motor learning-based upper-limb prosthesis control after home-based myoelectric training.
Keywords — feedback, electromyography, motor learning, hand prosthesis, abstract decoding, home training, long-term

1 Introduction

Muscles act as biological amplifiers of neural information transmitted by the central nervous system [1]. This neural information can be detected by electrodes placed on the surface of the skin that detect muscle activity in the form of electromyography (EMG) [1]. This muscle activity is used to control the most common form of active hand prostheses, myoelectric devices [2]. In these devices, user intent is typically estimated by extracting relevant information from the EMG and mapping it to a prosthesis output [3].

The mapping between muscle activity and prosthesis output depends on the control scheme. The most common approach for users with trans-radial (below-elbow) limb difference is “dual-site” control, which uses EMG signals from a residual muscle pair to provide bidirectional control of one degree of freedom (DoF). Controlling a multi-functional prosthesis using a single DoF at a time means dual-site control is typically sequential, thus limiting the approach [4]. Pattern recognition based on EMG offers an alternative to sequential control and is currently considered the probable successor to “dual-site” control [5]. With pattern recognition, it is possible to implement a physiologically congruent mapping of muscle activity to hand grasps [6,7]. The main advantage of this biomimetic approach is that it enables the functional use of the device to commence rapidly [8]. The drawbacks of pattern recognition are that it usually requires more than two EMG sensors, and it often involves regular recalibration because the approach is not flexible to the varying nature of EMG activity over time; this has made demonstrating robust control outside of the laboratory difficult [5,9,10].

1.1 Learning & Feedback

Myoelectric prosthesis control can be measured at various levels, from the decoding of intended grasps through to object manipulation. We use the term myoelectric control accuracy to describe the decoding of intended grasps, i.e. the probability that a prosthesis activates the grasp pattern intended by the user. Myoelectric control accuracy often improves when users interact with the system in real-time. Improvements in myoelectric control accuracy have been observed during the use of direct control [11], regression [12,13], and pattern recognition [5,11,14,15], and they have been documented in single sessions [12,13], multi-day studies [11,14–17], and multi-week studies [5,17]. None of these studies involved continual learning or time-adaptive algorithms. Although multiple independent factors will contribute to the gains observed, the only aspects that change over time in each study are the participants. Therefore, some form of human learning is the most reasonable explanation for the increase in performance. If learning occurs during prosthesis use, it follows that we should optimize the process, both during prosthesis training and prosthesis use.

This manuscript focuses on a novel approach for myoelectric training prior to receiving a prosthetic device. To improve training, we tested hypotheses based on the existing motor learning literature. It has been theorized that learning involves multiple processes acting on different timescales, sometimes in tension with one another [18,19]. One component of training motor tasks, the timing of feedback, has been shown to have different impacts on short-term and long-term learning [20–22]. Delaying feedback degrades the speed of skill acquisition and results in reduced after-effects following perturbation trials [22,23]. Because short-term studies commonly link an increased size of after-effects to learning, it has been suggested that delayed feedback is detrimental to learning [22,24]. In contrast, long-term retention studies have found that while skill acquisition is indeed slower with delayed feedback, a greater proportion of the performance observed during training is retained when feedback is withdrawn [25–28]. This is essential in the context of myoelectric training prior to prosthesis use, as the learned skill must be transferred to prosthetic control.
1.2 Abstract Decoding

Abstract decoding is a prosthesis control scheme that focuses on a user learning to generate novel patterns of muscle activity, such that distinct signals can be produced to control different prosthesis outputs [29–31]. This approach is one of numerous learning-based control schemes that have been proposed over the past decade that map EMG activity to prosthesis output non-biomimetically, i.e. the muscle activities used for control are arbitrary and do not align intuitively with motor intent [32–37]. A core premise of abstract decoding is that the plasticity inherent to the human nervous system can be exploited such that users can learn how to produce a) novel and distinct patterns of muscle activity and b) arbitrary mappings between said muscle activities and prosthesis outputs [29–31]. The general approach was defined in [30], and a later work demonstrated that amputees could learn to generate the novel muscle activities required [31]. Relative to other control schemes, abstract decoding offers reduced algorithmic complexity and sensor requirements; however, this comes at the cost of increased learning by the user. One aim of abstract decoding is to restore multiple hand grasps using two electrodes without cumbersome sequential switching [30,31].

To facilitate learning, abstract decoders present muscle activity in a non-representational multidimensional space [30]. In practice, this typically involves learning how to control a cursor using muscle activity in a center-out task based on visual feedback [29–31]. Presenting continuous feedback of muscle activity allows the motor system to generate an inverse map [38], which links motor output to arbitrary control variables for a given task space [39, 40]. The center-out task used to learn how to control the cursor acts as a visual representation of the task space. The hypothesis driving abstract decoding is that this space can be delineated into segments, each of which can be mapped to arbitrary prosthesis outputs, such as proportional digit control [29] or grasp selection [35, 37]. An underlying premise of abstract decoding is that the necessary muscle activity patterns can be learned prior to using a prosthesis [30]. For this premise to hold, long-term retention of motor skills is necessary because the user should be able to control the device without the precise feedback of muscle activity provided during training.

1.3 Retention

An idealized upper-limb prosthesis, one that is able to provide high fidelity concurrent closed-loop feedback of control signals and low latency prosthesis activation, would work with biofeedback and would not necessarily require retention of motor skill. Such systems do not currently exist. In terms of muscle activity, users only have access to noisy proprioceptive information [41], there is a delay between intention and any movement from the device [42], and the initial movements of the prosthesis are not necessarily informative. Therefore, it is crucial that users consistently reproduce the correct muscle activity for control in the absence of concurrent feedback. This necessitates that control tasks be learned, internalized and retained. In theory, this can occur either prior to prosthesis use or by training with a device; we focus on the former. Previous abstract decoding research has aligned with other learning-based control schemes in presenting concurrent visual feedback of the participant’s control signals [29–32, 34–37]; as outlined, this may not be optimal for ensuring retention of motor skills after myoelectric training.

Successful pre-device training in myoelectric prosthetics is often referred to as transfer [43]. Transfer denotes learning a skill in one domain, such as computer-based training, and applying it in a different domain (i.e., prosthesis control). Validating transfer in upper-limb prosthesis control is complex [43–45]. We believe that motor skill retention is a necessary precursor to transfer. Our prior research showed that an abstract decoding scheme can be learned in a closed-loop setting but not that learned activity is retained [30,31]. Therefore, this work aims to demonstrate the efficacy of delayed feedback on retention during pre-device training.
A laboratory- and a home-based study are presented. The purpose of the two studies was to investigate the retention of motor skills when using concurrent and delayed feedback paradigms. A delayed feedback protocol is introduced that postpones all visual feedback of the myoelectric interface until after the active control input has ceased. Two long-term experiments were performed. A laboratory-based study compared the use of concurrent \((n = 5)\) and delayed feedback \((n = 5)\) to assess the long-term stability of motor skill retention. Ten participants were trained using either concurrent or delayed feedback over four consecutive days with a follow-up probe in the third week. A home-based study established that the method introduced can be used to train myoelectric control outside the laboratory in appropriate time frames. We compared the real-time feedback and catch trial conditions used in our prior studies \([30,31]\) with new methods intended to periodically probe retention and assess long-term skill acquisition.

2 Methods

This section outlines the methods common to both experiments before outlining each experiment in detail. All participants recruited were able-bodied, free from neurological or motor disorders, and provided informed written consent. Ethical approval was granted by the local committee at Newcastle University (Ref: 20-DYS-050).

2.1 General

The methods used to estimate and calibrate muscle activity, the myoelectric task performed, and the feedback conditions used were common to both experiments.

2.1.1 Estimation of muscle activity

Muscle activity was estimated by computing the mean absolute value (MAV) over a 750 ms sliding window of rectified EMG data. Control signals were acquired at a consistent rate, which either exceeded or was close in value to the rate at which the display was updated. Therefore, changes in EMG activity were reflected in short time periods, equivalent to either one or two display frames. Muscle activity was smoothed over a window size found to balance responsive effector output against stable control during muscle contraction, as used in previous experiments \([30,31]\). The specific window sizes and control signal acquisition rates used are outlined in Sections 2.2 and 2.3.

2.1.2 Calibration

Instructions were provided on how to activate muscle groups via flexion and extension of the wrist. A calibration routine was then performed on the MAV filtered EMG data, \(y\). Participants were asked to perform dynamic arm movements while activity representative of baseline resting EMG, \(y_r\), was acquired. Participants then performed flexion and extension of the wrist to acquire data representative of comfortable contraction level, \(y_c\). An operator informed participants that contraction intensity should be limited such that it would be repeatable for long periods of time. In previous experiments this routine produced activity levels between 10 and 20% of the maximum voluntary contraction \([30,31]\). All subsequent experimental activity utilized a normalized muscle activation level, \(\tilde{y}\), calculated from \(y\) according to the following:

\[
\tilde{y} = (y - y_r) / (y_c - y_r).
\]

Feedback of the normalized activity was used to assess the participants’ ability to activate muscles for each control site independently. Visual inspection of raw and filtered EMG data was carried out prior to
each session. Once the normalization values for \( y_c \) were finalized, recalibration was strongly discouraged. Baseline EMG activity, \( y_r \), was only recalibrated at following sessions if changes in baseline noise affected control.

![Diagram](image)

Figure 1: The myoelectric control interface (MCI) task. (a) The two dimensional myoelectric controlled interface space. Cursor position is shown in green. (b) A representative cursor trajectory from basket to target. Thick cursor mark denotes the hold period. (c) Task timing structure for the Concurrent condition denoting cues, the move and hold periods. (d) Task timing structure for the Delayed condition denoting cues and the move, hold and playback periods. Dashed traces correspond to the ‘blind’ control input window. Solid traces refer to the playback of the cursor’s recorded path during the move and hold periods. The radius of the MCI on the display was \( \sim 64 \) cm. The MCI was viewed from a distance of 2 m.

### 2.1.3 Myoelectric control interface (MCI) task

The MCI task involved moving a cursor within the two-dimensional myoelectric control interface (MCI) outlined in Figure 1 [30, 31]. Cursor position on the interface was determined by \( \hat{y} \) for the two control sensors. The amplitude of activity in each muscle determines the cursor position along a single axis, such that the cursor would reach the upper-limit of an edge when \( \hat{y} = 1 \). The lower boundary of the targets is defined at \( \hat{y} = 0.3 \). The cursor is considered to be at rest when it is within the basket, i.e. \( \hat{y} < 0.2 \), as labelled in Figure 1b.

Participants were instructed to complete the task with their elbow flexed at 90°, whilst keeping the wrist in a neutral position. Trials commenced once participants relaxed their muscle activity such that the cursor was within the basket. An audible beep signaled the start of the trial, and then one of four targets was presented. Trials were \( \sim 1.5 \) s long and consisted of two periods of equal length, referred to as move and hold. Once a goal target was presented, the move period allowed the participant to react and begin moving the cursor out of the basket and toward the target. Participants were instructed that the aim was to keep the cursor within the target bounds during the hold period to maximize their score. A representative trial from a proficient user is shown in Figure 1b. At the end of a typical trial, a score was presented in the center of the screen.

The score refers to the proportion of the hold period during which the cursor was within or in contact
with the target. Because it is a ratio, the score is unitless and bound between 0 and 1.0. A score of 1.0 represents the cursor within or in contact with the target throughout the entire hold period, where timing is shown in Figure 1. A score of 0 represents the cursor never contacting the target during the hold period. In a collection of trials, targets were presented in pseudo-random order, such that each was experienced an equal number of times. The experiment was performed on a laptop computer (2.3 GHz i5-6200U CPU, 8 GB RAM, Lenovo, China). Real-time experimental software was implemented in Python using the AxoPy library [46].

2.1.4 Feedback conditions

The availability and timing of feedback during a trial were manipulated depending on the trial condition. The visual feedback conditions used in both experiments are detailed below:

Concurrent: During the trial period, the cursor position always reflected the normalized muscle activation levels, $\tilde{y}$, of the EMG channels used for control at that time frame. A score was presented at the end of the trial. The trial timing for the Concurrent condition is shown in Figure 1c.

Delayed: The cursor was invisible during the trial. At the end of the hold period, the trial cursor activity was played back to the participant at the same rate as it had occurred. Once the playback had concluded, a score was presented. Figure 1d shows the trial timing of the Delayed condition.

Catch: No feedback of the cursor’s movement was given; however, a score was presented after the trial.

Zero: No feedback of the cursor’s movement was given, and no score was presented after the trial.

2.1.5 Statistical analyses

All data were checked for normality via Shapiro–Wilk tests and visually inspected for confirmation. Where data were aggregated over days for statistical analyses, distributions did not present as normal. In addition, at the individual block level, the data did not tend toward being normally distributed. Therefore, we assumed that the data were not normally distributed and used non-parametric statistical tests for comparisons. All statistical significance calculations were performed using Mann-Whitney U tests. The Python library Scipy was used for calculations, using the recommended settings for all samples sizes [47]. Although we do not assume normality of data we present the mean unless otherwise stated. Our reasoning for this is as follows. At the block level, i.e. a collection of trials, the median scores in various conditions in both experiments rapidly attenuate to 1.0. In these circumstances presenting the median is therefore less visually informative of overall progression rates. At the group level, i.e. a collection of blocks, the median score would reflect the performance of a single participant at a time, which would also be less visually informative of overall rates.

2.2 Experiment 1: Laboratory-based comparison of feedback conditions

Experiment 1 was performed under laboratory conditions over four consecutive days. An additional follow-up session occurred in the third week, after a break of approximately 18 days.

2.2.1 Participants

Ten participants (two female, and eight male) were recruited. All participants were right-handed except for P7. The participants had either limited or no prior experience with the concurrent feedback condition of...
the myoelectric task. ‘Limited’ refers to short-term overall exposure to the task with no use over the prior 12 months.

2.2.2 Estimation of muscle activity

Eight channels of surface EMG signals were acquired using Trigno Quattro Sensors (Delsys Inc., Natick, MA, USA). Signals were acquired at 2000 Hz and band-pass filtered between 20 Hz and 450 Hz. The MAV filter for computing $y$ used a 750 ms window.

2.2.3 Calibration

Normalization values were permitted to change between the first two blocks of the familiarization period. After that, the calibration was deemed final. However, it was necessary to repeat the EMG calibration routine for one participant, P8, on the second day of training. All remaining participants used the normalization constants set on day one for the entire experiment.

To ensure consistent arm posture, the experiment used inertial measurement unit (IMU) data. Calibration of the IMU data was performed prior to each session. A plumb-line was used to ensure a consistent wrist angle across sessions. An image demonstrating use of the plumb-line is given in the Supplementary Material (Figure S1a). In the case of IMU signal drift, the calibration procedure was repeated within a session.

2.2.4 Recordings

Participants stood with their elbow flexed at 90° and their wrist in a neutral position, two meters viewing distance from a 55-inch screen (Philips Q-Line BDL5530QL) that displayed the MCI task. The physical distance from the base of the basket to MCI upper bound on the display was approximately 64 cm. Two Trigno Quattro Sensors (Delsys Inc., Natick, MA, USA) were used to acquire eight channels of EMG data from the right forearm. Muscle palpation was used to place two sensors targeting the extensor carpi radialis (ECR) and flexor carpi radialis (FCR), which were used to control the task. The remaining six sensors were equally distributed around the forearm. The base units of the two Trigno Quattro Sensors, which contain the ground electrode and the IMU, were placed on the upper arm and the distal end of the forearm, as shown in Supplementary Figure S1. Electromyography data from these additional sensors were included for future analysis. The unit housing the IMU sensor was placed distally on the forearm. Position data signals were acquired at 74 Hz. During the experiments, operators monitored participants for proper posture. If the participants’ arms moved out of position, a graphic appeared to assist re-positioning before the next trial commenced. To ensure consistency of recording sites over the days, sensor locations were marked on the arm using a pen. In the final session prior to the hiatus, electrode positions were marked on a Tubigrip which also included markings for physiological landmarks. The Tubigrip was used along with photographs to align the electrodes as closely as possible during the follow-up session. An image showing a participant performing the task may be found in the Supplementary Material (Figure S1b).

A total of 27,920 trials were collected. All EMG data were visually inspected for artefacts. Data from trials containing significant movement artefacts, electrical noise, or any other signs of non-physiological external influence, were rejected. The mean artefact rejection rate after manual visual inspection was 0.15% ± 0.3%.

2.2.5 Protocol

All participants underwent an initial familiarization period of four blocks, each block consisted of 80 trials of concurrent feedback, as shown in Figure 2a. The familiarization period was used to assess baseline performance on the control task. Calibration constants were adjusted if necessary within the first two
familiarization blocks. Based on their scores in the final familiarization block, participants were incrementally assigned to either the Concurrent or Delayed groups (Figure 2b). Allocation was performed such that differences in the median performance between the two groups were minimized at the time of the test, and the results are shown in Figure 2c. An alternative visual representation of Figure 2b, which shows the influence of individual participant allocation on group medians, may be found in the Supplementary Material (Figure S2). Depending on the assignment, either concurrent or delayed feedback would be experienced as a learning condition. All participants experienced catch and zero feedback trials throughout the experiment.

The experimental protocol contained two distinct block structures, acquisition and retention blocks, that comprised a collection of trials. Acquisition blocks refer to learning conditions (i.e., 80 trials of either concurrent or delayed feedback). In addition, eight catch trials were pseudo-randomly interleaved throughout each acquisition block, intended to periodically probe the state of the participants’ internal model. Whereas, retention blocks consisted of 40 consecutive zero feedback trials. As shown in Figure 2a, retention was assessed at the start and end of each day.
**Acquisition Score**: Score calculated over a block of 80 trials of either concurrent or delayed feedback, depending on the condition.

**Retention Score**: Score calculated over a block of 40 trials of zero feedback.

Prior to starting the experiment, an explanation of the muscle-cursor relationship was provided to expedite the exploration process. Participants were always informed of what to expect before experiencing a new block condition. Rests were permitted between blocks.

### 2.3 Experiment 2: Home-based training

Experiment 2 was carried out in the participants homes. Participants were provided with a network-enabled Arduino-based EMG acquisition device, a description of which may be found in [48].

#### 2.3.1 Participants

Four participants (two female, and two male) were recruited. All participants had previous experience with the concurrent feedback paradigm. There was no overlap in participants with Experiment 1.

#### 2.3.2 Estimation of muscle activity

Two channels of surface EMG were acquired using modified Gravity analog EMG sensors (OYMotion Technologies Co. Ltd, Shanghai, China) [48]. Each modified Gravity analog EMG sensor had an integrated ground electrode [48]. Signals were acquired at 500 Hz and band-pass filtered between 20 Hz and 150 Hz. The MAV filter for computing $y$ used a 760 ms window.

#### 2.3.3 Calibration

After setting the initial normalization values, it was necessary for participant P3 to recalibrate prior to day two. All other participants maintained the same calibration throughout the entire experiment.

#### 2.3.4 Recordings

The two Gravity EMG sensors (OYMotion Technologies Co. Ltd, Shanghai, China) were placed over the FCR and ECR muscle groups. Participants marked the location of sensors with a pen to ensure consistent positioning over the days. An EMG acquisition device was mounted on the forearm, and data were transmitted via Bluetooth to a computer [49]. Python-based experimental software was installed on laptop devices for use in the home or on the participants’ own devices. All data were automatically uploaded to a remote server for retrieval and analysis [48].

#### 2.3.5 Protocol

Participants carried out the experiment over five consecutive days and were given complete freedom regarding their training structure. Participants practiced at their own convenience and chose how many blocks to complete in a given session. Alternating between concurrent and delayed feedback blocks was permitted. The participants were informed that their goal was to maximize performance in the delayed feedback condition. At the end of each block, participants were presented with a plot of their progress, as well as their position on a leader board that ranked their performance against the other participants engaged in home-based training at the same time. Acquisition blocks consisted of 60 trials of either concurrent or delayed feedback trials, exclusively. Neither catch nor zero feedback trials were included in Experiment 2. Participants 1-4 completed 22, 31, 29, and 30 blocks, respectively, corresponding to a total of 6,720 trials.
Figure 3: Effect of feedback conditions on retention and acquisition. Alternating backgrounds show days. Axes (a) and (b) are aligned such that the points are chronological. (a) Group retention scores over training days and follow-up sessions. Envelopes represent the standard error of mean. (b) Group acquisition scores over training days and follow-up sessions. Points in (a) and (b) are the mean average of the participants' mean block score. (c-d) Concurrent and Delayed group scores over feedback conditions, respectively. Points in (c) and (d) are the collective mean over all corresponding trials completed on a given day. Asterisks indicate significant differences between Concurrent and Delayed training (Mann-Whitney U test, \( p < 0.05 \)).
3 Results

The results for each experiment are presented in separate sections.

3.1 Experiment 1

A comparison of average retention scores in the Concurrent and Delayed groups is shown in Figure 3a. As presented in Figure 2a, seven retention tests were carried out over the initial four training days, and a final two were completed in the follow-up session. Differences were calculated between the groups for each retention block. The means were calculated for day one retention score and a Mann-Whitney U test was calculated (Concurrent: 0.42 ± 0.11; Delayed: 0.40 ± 0.23; p = 0.68), inferring that the groups started at comparable levels. On day four, significant differences in performance were found in the first retention test (Concurrent: 0.32 ± 0.12; Delayed: 0.55 ± 0.15; p < 0.05) and the final retention test (Concurrent: 0.28 ± 0.13; Delayed: 0.62 ± 0.15; p < 0.05). There was no significant difference in initial retention performance during the follow-up session (Concurrent: 0.27 ± 0.12; Delayed: 0.48 ± 0.15; p = 0.1). However, after two refresher acquisition blocks, the Delayed group retention was significantly higher than the Concurrent group on the final block (Concurrent: 0.34 ± 0.13; Delayed: 0.63 ± 0.12; p < 0.05).

The average acquisition scores in the Concurrent and Delayed groups are shown in Figure 3b. Over the four-day training period the Concurrent group improved from an initial score of 0.75 ± 0.10 to 0.92 ± 0.03 in the final run. As expected, the Delayed group initially scored lower (0.45 ± 0.16). However, after four days of training, the Delayed group achieved an average final score of 0.70 ± 0.12. During the refresher acquisition blocks in the follow-up session, both groups were able to recover similar scores to those observed 18 days prior (Concurrent: 0.87 ± 0.10; Delayed: 0.68 ± 0.10). A figure illustrating the individual learning rate differences is shown in Figure S3 in the Supplementary Material.

The scores obtained over different trial conditions for each day are shown in Figures 3c and 3d. In the Concurrent group (Figure 3c) acquisition scores increased; however, no equivalent trend was observed in rate differences is shown in Figure S3 in the Supplementary Material.

Figure 4b compares the scores achieved over the peripheral and central targets during the final retention test during the follow-up session. No significant difference was found for peripheral target scores between the two groups (Concurrent: 0.53 ± 0.23; Delayed: 0.68 ± 0.16; p = 0.4). In contrast, a significant difference was found between the score achieved over central targets (Concurrent: 0.15 ± 0.11; Delayed: 0.58 ± 0.11; p < 0.05). In the absence of feedback, the Concurrent group tended to perform poorly over central targets, instead reaching the closest adjacent peripheral target, whereas the Delayed group was more able to perform the muscle co-contraction ratios necessary to reach the central targets. A plot comparing the average cursor angle error between the groups can be found in Figure S5 in the Supplementary Material.

3.2 Experiment 2

Data tracking the participants’ average scores over the five days of training are shown in Figure 5a. As days included different numbers of blocks, values presented are the mean and standard error of the mean. Comparisons were carried out using a Mann–Whitney U test with Bonferroni corrections at a significance level of p = 0.05. No trend of improvement was seen for P1 during training (Day 1: 0.92 ± 0.002; Day 5: 0.89 ± 0.009). Due to the number of blocks completed, significance values could not be calculated for Day 1 comparisons for Participant 1. No significant improvements were found between consecutive days for P2. However, a significant improvement was observed between the first and final days of training (Day
Figure 4: Retention differences between groups by target type. (a) Target type labels. (b) Final retention scores for peripheral and central interface targets for the Concurrent and Delayed participant groups. The points show the individual participants’ mean scores. Asterisks indicate a significant difference between Concurrent and Delayed training (Mann-Whitney U, p < 0.05).

1: 0.54 ± 0.061; Day 5: 0.91 ± 0.003). Participant 3 showed significant improvement after training (Day 1: 0.53 ± 0.033; Day 5: 0.83 ± 0.017). Additionally, significant differences were found on Days 2, 4 and 5 when compared with the preceding day. Finally, scores did not differ significantly between days for Participant 4 (Day 1: 0.83 ± 0.020; Day 5: 0.91 ± 0.013).

Modified confusion matrices, corresponding to the first and best performing delayed feedback blocks are shown in Figures 5b and 5c. The predicted target values for the confusion matrices were calculated offline. In each trial, the first target the cursor dwelled within consecutively for 240 ms was selected as the predicted target [49]. The confusion matrices presented have been modified to include a None class for the condition where the cursor did not dwell within any target for 240 ms consecutively within a trial (a null output). A value of 1.0 in the confusion matrices represents all predicted targets matching presented targets, i.e. when the MCI task presents said target the user can hold the cursor within the target for 240 ms without activating any other target. Highest overall prediction rates for each participant occurred on blocks 11, 23, 23, and 20 for participants 1-4, respectively. The confusion matrices for the first block (Figure 5b) show that initially more accurate predictions were achieved on the peripheral targets compared to the central targets. Figure 5c shows that after some training, the number of accurate predictions across targets generally increased, with central target predictions demonstrating the greatest improvement.
Figure 5: Overview of home-based training performance. (a-c) Referring to a column of plots, each row relates to the performance of a single participant. (a) Participants’ mean delayed feedback scores and cumulative blocks experienced over five days of training. Error bars correspond to the standard error of the mean. (b-c) Modified confusion matrices derived from the first and best delayed feedback blocks, respectively.
4 Discussion

Techniques from motor learning are increasingly being used in the context of upper-limb prosthetics [29,31, 33–37]. Historically, motor learning paradigms have provided concurrent feedback during training. While impressive performance can be achieved using feedback, this often has little meaning unless users have access to similar feedback when controlling their devices. This is problematic for motor learning-based myoelectric control schemes because real-world prosthesis users do not yet have feedback of their control input signals. Our results demonstrate that with appropriate training, it is possible to learn and to consistently reproduce distinct patterns of abstract muscle activity in the absence of concurrent feedback. As each distinct pattern can be mapped to a prosthesis output, four distinct patterns equate to a four class prosthesis control with pattern recognition. In addition, we demonstrated that the patterns of activity retained can be recalled after three weeks. This suggests that no algorithmic assistance or additional hardware above and beyond the clinical standard is necessary to restore four grasp classes using existing dual-site control devices.

Our results show that with appropriate training it is possible to learn and consistently reproduce distinct abstract muscle contractions in the absence of concurrent feedback. To demonstrate retention, we collected one of the largest online myoelectric control datasets with around 35,000 trials. Retention of skills can only be measured in the absence of the feedback provided during training. In Experiment 1, retention scores were calculated over individual blocks of 40 consecutive trials, where no feedback of the participants' control input or performance was provided. The results presented are clear and explainable. Figures 3a and 3b show that although lower scores are initially obtained with delayed feedback, the skills learned are retained multiple weeks later. Conversely, the higher performance gains observed with concurrent feedback dissipate during retention tests. Figure 4b explains why this difference in retention scores occurs. The contrast in retention performance can be attributed to different scores over the central targets. Only the delayed feedback group retained the motor skill of generating the muscle activation ratios necessary to reach the central targets. Similarly, Figure 5 shows that the improvements were largely due to participants increasing their scores over the central targets. Figure 5c shows the equivalent four-class confusion matrices, which reflect the upper bounds of what is possible with abstract decoding. The Supplementary Material contains a video that we believe is the first example of four grasp upper-limb prosthesis control based entirely on motor learning. In the video, a limb-intact participant accesses four prosthesis grasps without visual feedback, using an embedded implementation of the MCI task presented in this manuscript.

In the concurrent feedback condition, performance increased (Figure 3b) but learning did not (Figure 3a). Figure 3c shows that neither catch nor retention performance increased in the concurrent feedback condition. Our results clearly replicate the observation that frequent feedback can degrade learning. Multiple theories have attempted to explain this, and many of them have been summarized in [50]. Briefly, one view suggests that feedback not only has beneficial guiding effects on performance but also has a set of negative qualities that can be detrimental to learning. It is thought that feedback could encourage too much compensation during practice. Therefore, changing the input response too frequently may prevent a stable behavior from forming [50]. Additionally, as the user improves, a greater proportion of the learner’s total error can be attributed to neuromuscular noise. This noise is by nature un-correctable; therefore, any further compensations is maladaptive and continues to prevent consistent behavior [50]. Another theory suggests that frequent feedback may overwhelm attention and interrupts other learning processes (e.g., the development of internal error detection capabilities). While delayed feedback disrupts implicit learning, explicit learning is minimally affected [51]. Discovering a successful strategy early during practice may have played a key role in PI’s high accuracy on Day 1 of the home study. However, the underlying behavioral changes due to delayed feedback remain unclear, and an explanation is outside the scope of this work.

In Experiment 1, we compared the short-, intermediate-, and long-term stability of skill by probing the internal model at various stages throughout the experiment. Catch trials are regularly used in neuroscience to assess the state of internal models [52]. In previous research, we found that catch trial performance correlated with individual acquisition ability [30]. In this study, we compared catch trials with retention
be found in [58].

transfer of myoelectric control to prosthesis control using abstract decoding, and preliminary results may demonstrated solely by user control post-training [57]. The focus of our next work is to demonstrate the retention of myoelectric ability is necessary for transfer to prosthesis control. However, transfer cannot be demonstrated solely by user control post-training. As retention refers to the persistence of performance [56], we assume in Figure 3c, estimating retention requires consecutive reduced feedback trials to allow transient memory effects to dissipate.

One of the most unpredictable aspects of learning-based prosthesis control is individual differences in learning. The performance curves for individual participants are shown in Figure S3 (Supplementary Material). The curves show that it is not possible to predict how long it takes for participants to reach a fixed skill level; for example, in Experiment 1, Participant 8 took four days to reach the performance level that Participant 10 achieved on Day 1. Laboratory-based training must accommodate all learning rates which is sub-optimal for many other participants and the experimental operators. Home-based training has numerous advantages in this area [48]. We anticipated that, given time, most participants would be able to achieve high levels of performance. Therefore, we conducted a pilot study to investigate delayed feedback training at home. Unlike in the laboratory study, participants were given autonomy over how they conducted their training. While Experiment 1 focused on the underlying science, Experiment 2 bridged theoretical neuroscience with the practical implementation of our work. Modified confusion matrices were calculated post hoc to benchmark the levels of performance, enabling comparison to alternative techniques. Figure 5c shows that participants can reach four-class performance suitable for prosthesis control. Although the pilot was carried out over five days, this equates to only approximately 5 hours of training, with an average of 1680 trials per participant [48]. Our future experiments are likely to use a hybrid structure, in which participants train at home and perform rigorous testing in a controlled laboratory environment.

The peripheral targets of the interface only require independent muscle activation and are therefore easier to reach. This performance pattern matches previous results [30] and is repeated at the start of this experiment, see Figure S5 (Supplementary Material). Figure 4 shows that delayed feedback facilitates learning of the ratios of muscle activity necessary to reach the central targets, while concurrent feedback does not. This suggests that concurrent feedback encourages iterative ad-hoc correction within the trial period. This would be indicative of adaptation, in the prosthesis definition of the word, during extremely short intervals. Little is understood regarding motor-learning without concurrent feedback [51]. We have demonstrated that the ability to produce four distinct muscle patterns, equating to direct access to four prosthesis grasps, can be retained without requiring any algorithmic methods. It is unknown whether additional muscle contraction ratios could be retained from the two control sites. However, additional ratios are likely to be limited by inherent noise in the system rather than the central nervous system's capability to learn.

From the literature, we infer that the average number of grasps used by prosthesis users is four [53–55]. The abstract decoding approach can work with two electrodes, matching the current clinical standard while mitigating the drawback of cumbersome mode switching. Based on the data presented here and in the Supplementary Material, we believe that it is feasible to restore four grasps using abstract decoding. Our work demonstrates retention, and the video in the Supplementary Material demonstrates abstract myoelectric control after training. As retention refers to the persistence of performance [56], we assume retention of myoelectric ability is necessary for transfer to prosthesis control. However, transfer cannot be demonstrated solely by user control post-training [57]. The focus of our next work is to demonstrate the transfer of myoelectric control to prosthesis control using abstract decoding, and preliminary results may be found in [58].

One limitation of this study is the use of limb-intact participants. However, we infer that the underlying
learning mechanisms in people with limb-difference are equivalent and are not impacted by amputation [31, 59–61]. While task acquisition can be slower in amputees, this is likely due to increased sensitivity to fatigue, which leads to reduced training time [31]. Our previous work showed that amputees are capable of generating the muscle activity required in abstract decoding tasks [31]. Our study is also limited by small sample sizes. Experiment 1 was logistically challenging for a laboratory-based study and involves significantly more trials than is common in myoelectric control research. We therefore used the minimum number of participants necessary for statistical power. Experiment 2 was introduced to demonstrate the feasibility of using delayed feedback training in a more efficient manner in the home. Based on the first four participants tested, we concluded that the approach was feasible and stopped testing. The present work demonstrates that individuals can learn to produce the necessary muscle activity in the absence of feedback. To achieve this, our study design compared concurrent and delayed feedback as discrete conditions. As such, the fixed-length familiarization period of concurrent feedback was not optimal for learning across individuals. In applied conditions, participants transition at bespoke moments. The relationship between concurrent and delayed feedback may explain the differences in results obtained in Experiments 1 and 2. Participants in Experiment 1 either had limited or no prior experience with the task. Participants in Experiment 2 trained with delayed feedback after experiencing concurrent feedback. Although solely intended to demonstrate practical pre-device training, it is likely that Experiment 2 bore greater similarity to optimal late-stage biofeedback training than Experiment 1 [62].

Acknowledgements

This work has been supported by the Engineering and Physical Sciences Research Council (EPSRC) via grants EP/R511584/1 and EP/R004242/2 and the National Institute for Health Research (NIHR) / Devices for Dignity (D4D) Starworks Proof of Concept Funding (project: STWK-006).

References


