The Effect of Feedback During Training Sessions on Learning Pattern-Recognition-Based Prosthesis Control

Morten B. Kristoffersen, Andreas W. Franzke, Corry K. van der Sluis, Alessio Murgia, *Member, IEEE*, and Raoul M. Bongers

***Abstract*— Human-machine interfaces have not yet advanced to enable intuitive control of multiple degrees of freedom as offered by modern myoelectric prosthetic hands. Pattern Recognition (PR) control has been proposed to make human-machine interfaces in myoelectric prosthetic hands more intuitive, but it requires the user to generate high-quality, i.e., consistent and separable, electromyogram (EMG) patterns. To generate such patterns, user training is required and has shown promising results. However, how different levels of feedback affect effectivity in training differently, has not been established yet. Furthermore, a correlation between qualities of the EMG patterns (the focus of training) and user performance has not been shown yet. In this study, 37 able-bodied participants (mean age 21 years, 19 males) were recruited and trained PR control over five days. Three levels of feedback were tested for their effectiveness: no external feedback, visual feedback and visual feedback with coaching. Training resulted in improved performance from pre- to post-test with no interaction effect of feedback. Feedback did however affect the quality of the EMG patterns where people who did not receive external feedback generated higher amplitude patterns. A weak correlation was found between a principal component, composed of EMG amplitude and pattern variability, and performance. Our results show that training is highly effective in improving PR control regardless of feedback and that none of the quality metrics correlate with performance. We discuss how different levels of feedback can be leveraged to improve PR control training**

***Index Terms*— Electromyography, motor learning, myoelectric control, prosthesis, pattern recognition.**

I. Introduction

MYOELECTRIC prosthetic hands have evolved from one degree of freedom (DoF) grippers to multi-articulated hands with multiple DoFs. However, the commercially available human-machine interface that should allow control of those multiple DoFs has not yet advanced to make the control process intuitive [1]. In particular, the Direct Control (DC) scheme used by these hands only allows for control of one DoF at a time. Moreover, switching between DoFs is done sequentially using electromyography (EMG) trigger signals generated using unintuitive muscle contractions. Pattern-Recognition (PR) control has been proposed as an alternative, since the PR algorithm can learn to predict movement intent using EMG patterns generated from the corresponding (phantom) movement. This would allow for almost instantaneous switching between grips which could lead to faster, more fluent and more intuitive control [2]–[4].

A prerequisite for good PR control is that EMG patterns have specific qualities. In the context of PR this means that EMG patterns are sufficiently distinct, not too variable and can be generated with high reliability [5]–[7]. However, since the EMG patterns are measured at the surface of the skin, the quality of the EMG patterns is negatively affected by non-stationarity effects such as sweat and electrode shifts [2], [8]. In addition, movement repetition variability can also negatively affect specific qualities of the EMG patterns so that they differ between repetitions, since humans cannot naturally generate EMG with such high precision [9], [10]. While non-stationarities can be reduced by improving algorithms' classification accuracy and reducing noise from data acquisition, the quality of the generated EMG patterns can only be improved by training the user.

Although user training for PR control has been studied previously [6], [11], results are inconclusive as studies differ in methods, results and lack statistical power due to small sample sizes [5], [6], [11], [12]. All the aforementioned studies do however report significant improvements after training and most attribute the improvements to changes in a set of representative EMG features [5], [6], [12]. This set spans a feature space and changes in said feature space have been hypothesized to correlate with performance improvements.

When analyzing the feature space, metrics such as the distance between EMG patterns of different movements (separability), the distance between EMG patterns belonging to the same movement (consistency), and the size of the EMG patterns (variability) are used to assess the quality of the EMG

Manuscript received January 24, 2019; revised April 17, 2019 and June 12, 2019; accepted July 12, 2019. Date of publication August 20, 2019; date of current version October 8, 2019. This project (INPUT) has received funding from the European Unions's Horizon 2020 Research and Innovation Programme under Grant 687795. *(Corresponding author: Morten B. Kristoffersen.)*

M. B. Kristoffersen, A. W. Franzke, and C. K. van der Sluis are with the University of Groningen, University Medical Center Groningen, Department of Rehabilitation Medicine, 9713 GZ Groningen, The Netherlands (e-mail: m.b.kristoffersen@umcg.nl; a.w.franzke@umcg.nl; c.k.van.der.sluis@umcg.nl).

A. Murgia and R. M. Bongers are with the University of Groningen, University Medical Center Groningen, Center for Human Movement Sciences, 9713 AV Groningen, The Netherlands (e-mail: a.murgia@umcg.nl; r.m.bongers@umcg.nl).

Digital Object Identifier 10.1109/TNSRE.2019.2929917

patterns [3], [5], [6], [12]. In studies on user training for PR control, EMG pattern quality improved, albeit in different ways. The level of feedback given during training differed among the studies which arguably played a role in those differences. The feedback ranged from no feedback [12], to visual feedback [5], to visual feedback with individualized coaching [6]. It is currently unknown which aspects of EMG pattern quality are the most important for good prosthetic performance, as no clear correlation between any metrics derived from the feature space and PR control performance has been shown. To develop an effective training scheme for PR control it is important to know which aspects of EMG pattern quality to target and which level of feedback improves this most [5].

The main goal of this study is to determine the effect of the level of feedback on EMG pattern quality and how this influences PR control performance. The primary outcome variables used are the test scores (number of completed movements and online accuracy) measured after training, which describe PR control performance. The secondary outcome variables are a selection of metrics describing EMG pattern separability, consistency and variability, which overall describe EMG pattern quality. We hypothesize that (1) higher levels of feedback result in higher test scores and (2) that higher levels of feedback results in larger changes in quality of the EMG patterns and (3) that a strong correlation exists between the primary and secondary outcome variables.

II. Methods

A. Participants

Participants were eligible if they were able bodied with no physical limitations. The study took place in the laboratories of the Department of Human Movement Sciences at the University of Groningen, the Netherlands. Four junior researchers recruited the participants using their social networks and by advertising the study on student forums. Prior to participation all participants were informed about the experiment and signed a letter of informed consent. This study was approved by the local ethics committee (ECB/2017.01.12_1).

B. Materials

Participants were seated at a table with their non-dominant hand resting at the edge of that table and the elbow resting on the chairs' arm rest. Hand dominancy was determined using an adapted version of the Edinburgh inventory [13], [14].

A brace (Medical Specialties Wrist Lacer) was attached to the distal part of the non-dominant side, restricting the movements of the wrist and thumb. By restricting the wrist and thumb, participants produce isometric muscle contractions similar to those of a person with an upper-limb defect. [15], [16]. We restricted the hand, to improve the transferability of our results to individuals with an upper limb defect.

Eight active bi-polar electrodes (Otto Bock 13E200=50AC) were attached equidistantly around the proximal side of the participants' non-dominant forearm. EMGs were sampled at 1 kHz and transmitted wirelessly to a laptop computer

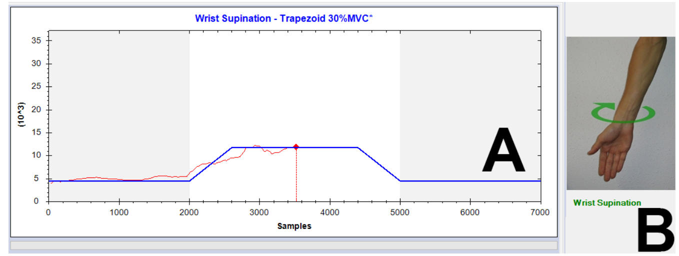

Fig. 1. Screen feedback given during system training in the pre- and post-test. A: Visual cue, the height of the moving red dot was determined by averaging the root mean square of the EMG over all electrodes. The participant was asked to match the height so that the red dot followed the blue trapezoid shape. B: picture of the requested movement. MVC=Maximal Voluntary Contraction.

(Dell XPS 15 9550). Participants started training minutes after putting on the electrodes and noise was assessed before training by visual perusal of the EMG. Four time-domain features of the EMG (mean absolute value, waveform length, zero crossings and slope changes) were extracted from 128 ms windows with 32 ms overlap [17]. The features were classified into one of 8 classes (7 movements + rest) using a Linear Discriminant Analysis (LDA) classifier and mapped to a movement and velocity.

C. Procedure

The experiment was conducted over five days, and consisted of daily sessions. EMG production during each session was limited to 30 minutes to avoid fatigue. On day 1 a pre-test was conducted followed by a 20-minute user training session. On days 2-4 the participants followed a 30-minute user training session. On day 5 the participants followed a 20-minute user training session and subsequently performed the post-test. Pre- and post-test lasted ~10 minutes. By changing the order between test and training on days 1 and 5, training time was maximized. Training, as well as the pre- and post-tests, consisted mainly of system training and Motion test procedures. Total session time was between 45 and 60 minutes.

1) Pre-Test and Post-Test: The pre-test and post-test were identical and consisted of two parts: System training and Motion test [18].

a) System training: In PR control, system training is the procedure in which the user supplies training data to construct the feature space on which the PR algorithm is trained. In the current experiment system training was done with seven movements: supination, pronation, wrist flexion, wrist extension, hand open, key grip and fine pinch. Each of the seven movements was executed four times for three seconds: once at 100% maximum voluntary contraction (MVC), which was only used to calibrate the height of the visual cue on the screen (see Figure 1). Participants had to follow this cue when they performed subsequently the movement at 30%, 60% and 90% MVC respectively [17]. Each movement was demonstrated by the experimenter before and during system training. The movement executions at 30%, 60% and 90% MVC were used to train the classifier.

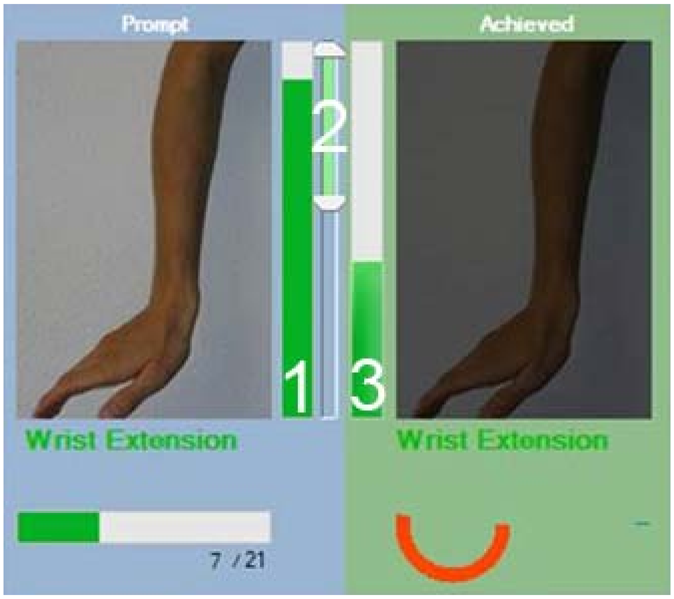

Fig. 2. Screen feedback given during the Motion test in the pre- and post-test. Left: Picture of the requested movement. The green bar (1) denotes the requested EMG amplitude and the smaller green bar with the handles (2) denotes the margin that the EMG amplitude from (1) should be within. Right: Picture of the predicted movement. The green (3) bar denotes the measured EMG amplitude. The red circle denotes the remaining time (in this instance, almost half the time has passed, as shown by the semi-circle). When the movement was matched, a green circle started to form next to the red circle. The green circle was full when the movement was successfully completed.

b) Motion test: In the Motion test participants performed the same movements as in the system training, but now the PR algorithm was used to classify the movement performed by the participant. This was done using the feature space constructed with the data from the system training. In the Motion test, participants performed 21 movement trials; all movements from the system training at 30% (±15%), 60% (±20%) and 90% (±30%) MVC respectively, presented in random order. While performing the movement, participants saw an image of the requested movement, the predicted movement and a timer on a screen, as shown in Figure 2. A movement trial was considered successful if the participant could perform the requested movement in such a way that the PR algorithm classified it for a minimum of two seconds within a three second timeout at the requested MVC level. The Motion test is comparable to the test and training used in the study of Bunderson and Kuiken [5].

D. Training Program

Identical to the pre- and post-test the user training sessions consisted of the system training and the Motion test, albeit with different feedback depending on the group to which the participant was assigned: no feedback, visual feedback, and visual+coaching feedback (Experimental groups, see section E). System training and Motion test were chosen for training since proper execution of both is necessary to achieve good prosthesis control. Both were conducted as training, twice on day 1 and 5 (after and before pre-test and post-test, respectively) and thrice on days 2-4. Participants were assigned to a specific group based on the inclusion date, with 1/3 of consecutive participants in each group.

E. Experimental Groups

All the groups received feedback about their performance in the pre-test and in the post-test; after the Motion test the number of correct movements and the online accuracy (Performance metrics, see section F1) were shown on the screen. The groups differed with respect to the level of feedback they received during training, as described below.

1) No Feedback [NF]: The NF group followed the training program, but without visual screen feedback. This meant they did not receive any feedback on how to perform the movements nor any feedback on their performance during training. More specifically, they did not see the content presented in Figures 1 and 2 during the user training sessions but only during the pre- and post-test. Participants in this group did not receive any visual feedback and relied on the cues provided by the experimenter, who performed the movements and told the participants when to start and stop executing the movements.

2) Visual Feedback [VF]: The VF group followed the training program while receiving visual screen feedback (see Figures 1 and 2) as in the pre-test and post-test.

3) Visual+Coaching Feedback [VCF]: The VCF group received the same visual feedback as the VF group while in addition they received coaching from the experimenter after each Motion test. The coaching was based on different metrics (see section F2a1, F2b1 and F2c1) of EMG pattern consistency, separability and variability. Before coaching, the experimenter reviewed a plot for each metric. To train consistency, the experimenter would instruct the participant to focus on the execution of the movement with the worst reproducibility (highest within-class distance). To train separability, the experimenter would instruct the participant to adapt the least separable pair of movements so as to make them more separate in terms of EMG patterns, following the general principle used by Powell and Thakor [19], see Figure 3. The experimenter also used a spider plot showing the root mean square of the EMG patterns to explain to the participant how the movement adaptations would affect the patterns and make them more separable, see Figure 3 [20]. To reduce variability, if the movement with the highest variability was also in the pair of the least separable movements, the experimenter would instruct the participant to focus on the execution of the movement with the highest variability (highest mean semi-principal axis) to make it less variable.

F. Metrics

1) Performance Metrics: In the Motion test two feedback metrics were computed, number of completed movements and online accuracy. These two feedback metrics were the primary outcome variables.

a) Number of completed movements: The number of completed movements was the total number of successful movement trials (maximum 21).

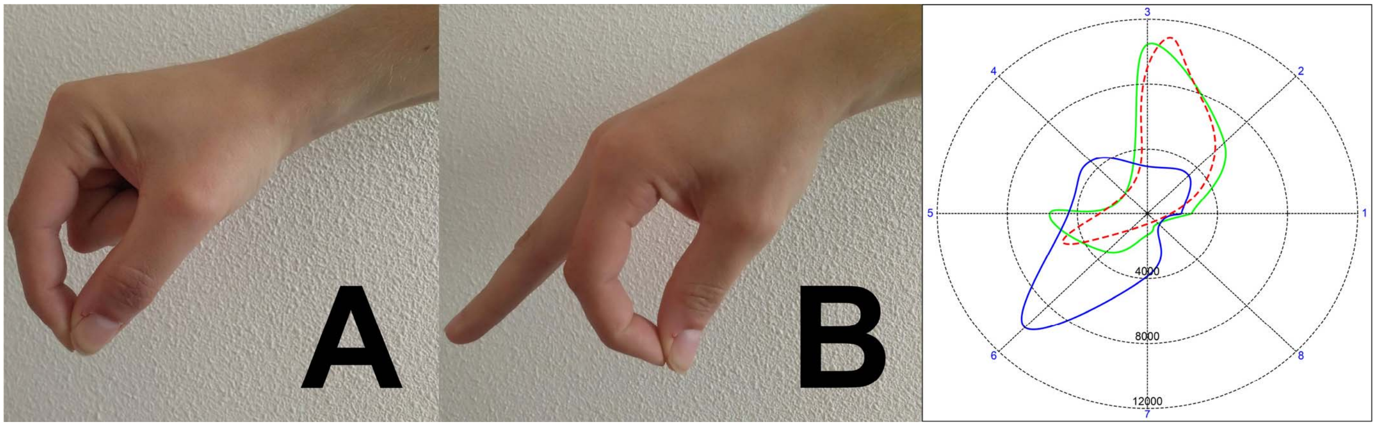

Fig. 3. Example of movement adaptation and feedback given to VCF group. A: fine pinch grip with the non-essential fingers flexed. B: fine pinch grip with the non-essential fingers extended. Right: Example of spider plot with EMG patterns from three movements plotted. The spider plot shows the root mean square of the EMG pattern for each channel and serves as a simplified representation of the feature space with each movement (EMG pattern) represented by one colored shape. Overlap can be visually perused and the experimenter can guide the participant to minimize it. Variations A and B of fine pinch are represented by the blue and the green shape respectively. For clarity we added a red dotted shape, representing an example EMG pattern that has considerable overlap with variation B of fine pinch. To show an example of an extreme case, the red shape was fabricated and is not based on EMG of a real grip. In this case, the experimenter would ask the participant to execute the fine pinch such as in variation A, to find the variation that causes the least overlap with the EMG pattern represented by the red dotted shape.

b) Online accuracy: The online accuracy denoted the percentage of time participants conducted the movement in such a way that the classifier could correctly predict it independently of whether the movement trial was successful or not. The participant could always see the current prediction as depicted on Figure 2.

2) EMG Pattern Quality Metrics: To determine the specific EMG pattern qualities related to PR control several metrics describing EMG pattern consistency, separability and variability were computed. These metrics were the secondary outcome variables.

a) Metric of EMG pattern consistency:

(1) Within-class Distance (WD)
The WD is a consistency metric of EMG patterns over repetitions of the same movement.

$$WD_j = \sum_{k=1}^{3} \frac{dist_{kj}^{rj} * dist_{rj}^{kj}}{dist_{kj}^{rj} + dist_{rj}^{kj}} \tag{1}$$

Where $dist_{kj}^{rj}$ and $dist_{rj}^{kj}$ are half the Mahalanobis distances in feature space between the EMG patterns of repetitions r and k of movement j and between repetitions k and r of movement j respectively:

$$dist_{kj}^{rj} = \frac{1}{2}\sqrt{\left(\mu_{Trj} - \mu_{Tkj}\right)^T * S_{Trj}^{-1} * \left(\mu_{Trj} - \mu_{Tkj}\right)} \tag{2}$$

$$dist_{rj}^{kj} = \frac{1}{2}\sqrt{\left(\mu_{Tkj} - \mu_{Trj}\right)^T * S_{Tkj}^{-1} * \left(\mu_{Tkj} - \mu_{Trj}\right)} \tag{3}$$

where μ_{Trj} and μ_{Tkj} denote the feature vectors containing the four features calculated from the system training data from repetition r and k respectively. $\mathrm{S_{Trj}}$ and $\mathrm{S_{Tkj}}$ are the covariances of the training data from repetition r and k respectively. When computing the total WD for a participant and not just one movement, we took the mean WD over all movements

$$WD_{total} = \frac{1}{7}\sum_{j=1}^{7}\left(\sum_{k=1}^{3} \frac{dist_{kj}^{rj} * dist_{rj}^{kj}}{dist_{kj}^{rj} + dist_{rj}^{kj}}\right) \tag{4}$$

b) Metrics of EMG pattern separability:

(1) Inter-class Distance Nearest Neighbor (IDNN)
The IDNN is based on the Separability Index (SI) of Bunderson and Kuiken [5] and is a separability metric of EMG patterns over different movements. We defined IDNN as

$$IDNN_i = \min_{i=1,\ldots j-1, j+1,\ldots 7} \frac{dist_j^i * dist_i^j}{dist_j^i + dist_i^j} \tag{5}$$

Where $dist_j^i$ and $dist_i^j$ is half the Mahalanobis distance in feature space between movements i and j and between movements j and i of movement j respectively

$$dist_j^i = \frac{1}{2}\sqrt{\left(\mu_{Ti} - \mu_{Tj}\right)^T * S_{Ti}^{-1} * \left(\mu_{Ti} - \mu_{Tj}\right)} \tag{6}$$

$$dist_i^j = \frac{1}{2}\sqrt{\left(\mu_{Tj} - \mu_{Ti}\right)^T * S_{Tj}^{-1} * \left(\mu_{Tj} - \mu_{Ti}\right)} \tag{7}$$

where μ_{Ti} and μ_{Tj} denote the feature vectors containing the four features calculated from the training data from movement i and j, respectively. $\mathrm{S_{Ti}}$ and $\mathrm{S_{Tj}}$ are the covariance of the training data from movement i and j, respectively. This distance metric is the same as the one for WD (see equations 2 and 3) except that instead of measuring distances between clusters from the same movement the distance is measured between clusters of different movements. The IDNN differs from the SI in that the Mahalanobis distance between clusters are measured in both directions instead of only one, meaning that the covariance of both clusters is considered. The SI differs depending on which cluster is measured from ($dist_j^i \neq dist_i^j$) [21], [22] whereas the IDNN between two clusters are always the same. When computing the total IDNN for a participant and not just one movement, we took the mean IDNN over all movements

$$IDNN_{total} = \frac{1}{7}\sum_{j=1}^{7}\left(\min_{i=1,\ldots j-1,j+1,\ldots 7}\frac{dist_j^i * dist_i^j}{dist_j^i + dist_i^j}\right) \tag{8}$$

(2) Inter-class Distance all Neighbors (IDAN)
The IDAN metric is similar to IDNN, with the difference that instead of measuring the distance to the nearest neighbor only, distances to all neighbors are calculated. Using the same notation as for IDNN (equation 5-8), we defined IDAN as

$$IDAN = \frac{1}{7}\sum_{j=1}^{7}\left(\frac{dist_j^i * dist_i^j}{dist_j^i + dist_i^j}\right) \tag{9}$$

(3) Most Separable Dimension (MSD)
The MSD metric is similar to IDNN, with the difference that instead of measuring the distance using the entire feature vector (i.e. all 32 dimensions) only the dimension in which the separability to the other movements is largest is used to determine the nearest neighbor. Using the same notation as for IDNN (equations 5-8), we defined MSD as

$$MSD = \frac{1}{7}\sum_{m=1}^{7}\max_{d=1,\ldots,32}\left(\min_{1=1,\ldots j-1,+1,\ldots 7}\frac{dist_j^i * dist_i^j}{dist_j^i + dist_i^j}\right) \tag{10}$$

where the feature vectors of dist (μ_{Ti} and μ_{Tj} in equation 6-7) are calculated for each of the 32 dimensions ($\mathrm{d} = 1, \ldots, 32$) and the dimension where the distance is the largest (max) is then used to calculate MSD.

c) Metric of EMG pattern variability:

(1) Mean Semi-principal Axis (MSA)
The MSA is unaltered from the formulation given by Bunderson and Kuiken [5] except that for feedback we only calculated the MSA for a single movement as opposed to the average of all movements as done by Bunderson and Kuiken. The MSA is a metric for how variable a movement of a participant is and is formulated as

$$MSA_k = \left(\prod_{k=1}^{16} a_k\right)^{\frac{1}{16}} \tag{11}$$

where a_k is the geometric mean of the semi-principal axes of the hyperellipsoids. When computing the total MSA for a participant and not just one movement, we took the mean MSA over all movements

$$MSA_{total} = \frac{1}{7}\sum_{j=1}^{7}\left(\left(\prod_{k=1}^{16} a_{jk}\right)^{\frac{1}{16}}\right) \tag{12}$$

d) Metric of EMG pattern amplitude:

(1) Mean Absolute Value (MAV)
The MAV metric was used to determine the EMG amplitude as EMG amplitude correlates with force. This metric was used in conjunction with the other metrics to determine if a change was the result of increasing force or of change in dimensionality

$$MAV = \frac{1}{7}\sum_{m=1}^{7}|EMG_m| \tag{13}$$

where EMG_m is the EMG of movement m of the training data.

G. Data Analysis

Several of the metrics used in the analyses were the same as those used to compute feedback to the participants in the VCF group (WD, IDNN, MSA). For the analyses the metrics were averaged over all movements as was described in section F2. All dependent variables were tested for normality in separate Kolmogorov-Smirnov tests.

To establish the effect of feedback on movement performance and EMG pattern quality, training outcome results and EMG patterns from the pre- and post-test were analyzed in separate ANOVAs. These analyses were conducted using two-way repeated-measures ANOVAs with test outcome (pre-test vs post-test) as the within-subjects factor and group (feedback level) as the between-subjects factor, using each of the primary and secondary outcome variables as the dependent variable in separate ANOVAs. Effect sizes were calculated using generalized eta-squared statistics [23]. Post-hoc testing was done using T-test with Bonferroni correction.

A Principal Component Analysis (PCA) [24] was additionally conducted to determine if any principal components (PCs) derived from the secondary outcome variables were correlated with any of the primary outcome variables. The PCA was performed on the percentage of change from pre- to post-test for each participant (37 observations) with variables represented by one of the EMG metrics described above (6 metrics). Used in this manner, PCA can reveal if a combination of metrics can explain the effect of feedback on training. The PCs thus obtained, which are as many as the metrics, represent the normalized linear combination of the metrics that explain a percentage of the variance in the percentage of change. This is done so that the first PC has the largest variance, the second PC has the second-largest variance and so on. The loading vectors represent the directions of variability of the data in the variable space, with the first loading vector as the direction along which the data vary the most and so on. We analyzed the set of PCs that together explained 90% of the variance. The PCs between groups were compared using ANOVAs with group as between subject factor using the PCs as dependent variables. Post-hoc testing was done using the T-test with Bonferroni correction.

To determine if there was a correlation between pattern change and training outcome, we calculated Pearson's correlation coefficient between the primary and secondary outcome variables. Pearson's correlation coefficient was also calculated between the primary outcome variables and the PCs.

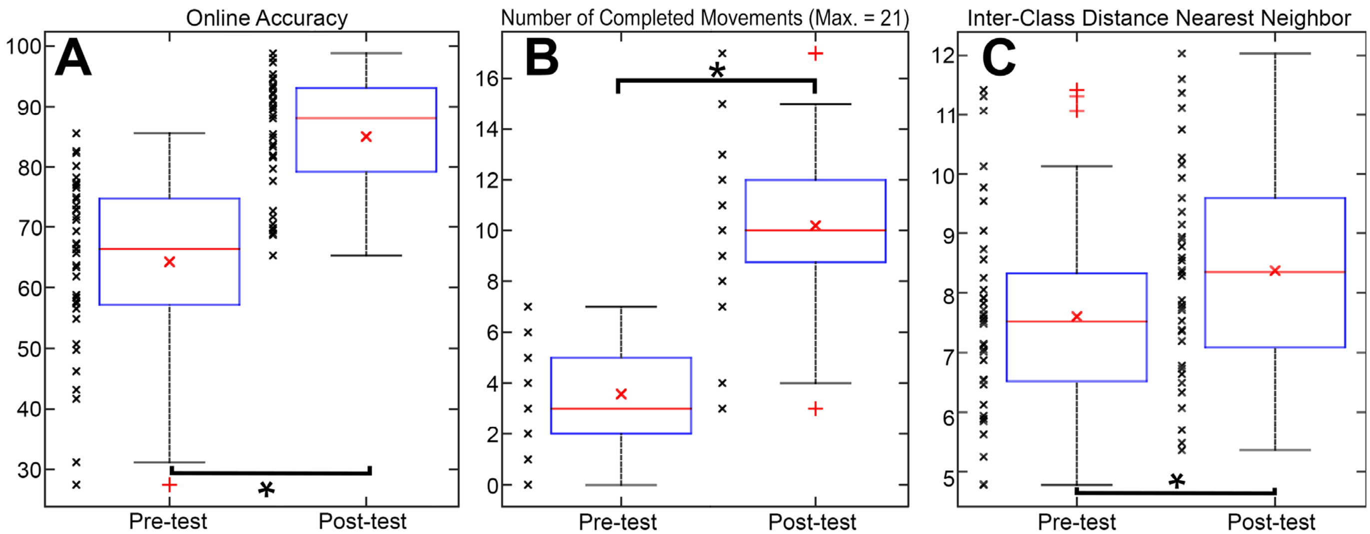

Fig. 4. Boxplots showing (A) online accuracy (B) number of completed movements and (C) Interclass Distance Nearest Neighbor (IDNN) from pre- to post-test. The red crosses denote the mean and the red line in each plot is the median. Outliers, marked with plus signs, are defined as points that lie outside 1.5 the interquartile range. Individual data points are marked with black crosses.

Results are reported using mean ± standard error of the mean. The PCA analysis was conducted using Matlab R2016a and all other analysis was conducted using IBM SPSS Statistics version 23.

III. Results

Thirty-seven participants were included in the experiment (N=37, female=18; mean age 21.6 ± 2.1 years). There were 13 participants assigned to the NF group, 12 to the VF group and 12 to the VCF group.

The Kolmogorov-Smirnov tests showed that all dependent variables followed a normal distribution.

A. Effect of Test and Feedback on Primary Outcome Variables

The analysis on online accuracy revealed a significant effect of test (F(1,34) = 48.41; p < .0001; η_G^2 =.0009) with an increase between pre- and post-test (64% ± 2.3%; 85% ± 1.6%, Figure 4a). The analysis on the number of completed movements also revealed a significant effect of test (F(1,34) = 175.121; p < .001; η_G^2 = .66) with an increase between pre- and post-test (3.56 ± .29; 10.18 ± .47, Figure 4b)[1]. In both analyses none of the effects involving the three feedback groups were significant.

B. Effect of Test and Feedback on Secondary Outcome Variables

1) Metric of EMG Pattern Consistency: No significant effects were found for WD.

[1]To assess if the increase in the primary outcome variables was due to training or was the result of a random effect, we performed a repeated measures ANOVA on the results from the intermediate Motion tests (first Motion test after pre-test, last Motion test on days 1-4 and the last Motion test before the post-test) as within-subject factor and Feedback group as between-subject factor. The ANOVA showed only a significant main effect (p < .0001) of the tests, thus we concluded that the effect is due to training.

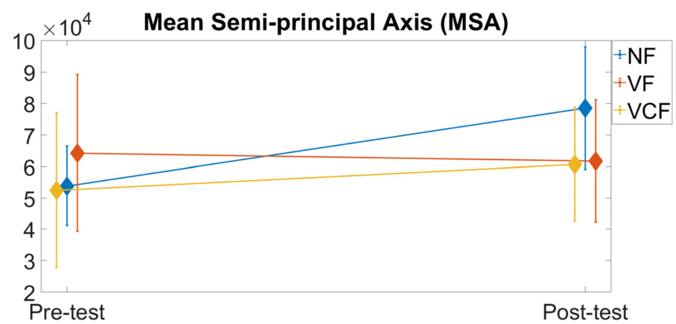

Fig. 5. Mean semi-principal axis for all groups from pre- to post-test. Error bars shows the standard deviation.

2) Metrics of EMG Pattern Separability: IDNN showed a significant effect of test (F(1,34) = 8.326; p = .007; η_G^2 = .05) with an increase from pre- to post-test (7.6 ± .27; 8.37 ± .29, Figure 4c). None of the other effects were significant. No significant effects were found for IDAN and MSD.

3) Metric of EMG Pattern Variability: MSA revealed a significant effect of test (F(1,34) = 12.174; p = .001; η_G^2 = .064) with an increase between pre-test and post-test (56,764 ± 3,521; 67,297 ± 3,333). This effect of test interacted with that of group (F(2,34) = 7.53; p = .002; η_G^2 = .078). Post-hoc analysis using Bonferroni correction revealed that participants in group NF increased MSA between pre-test and post-test (p < .0001) while participants in groups VF and VCF did not change (p = .584 and .199 respectively), see Figure 5.

4) Metric of EMG Pattern Amplitude: MAV revealed a significant effect of test (F(1,34) = 9.983; p = .003; η_G^2 = .06) with an increase between pre-test and post-test (1,101 ± 70; 1,306 ± 70). This effect of test interacted with that of group (F(2,34) = 7.842; p = .002; η_G^2 = .09). Post-hoc analysis using Bonferroni correction revealed that participants in group NF increased MAV between pre-test and post-test (p < .0001) while participants in groups VF and VCF did not change (p = .486 and .372 respectively), see Figure 6.

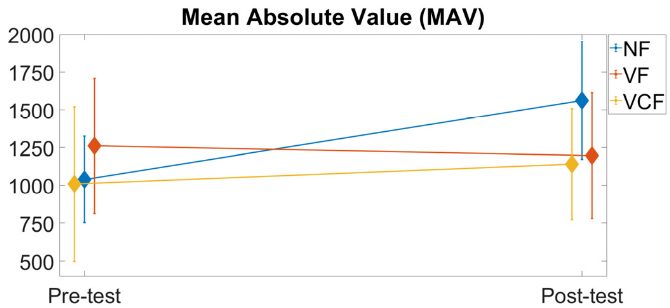

Fig. 6. Mean Absolute Value for all groups from pre- to post-test. Error bars shows the standard deviation.

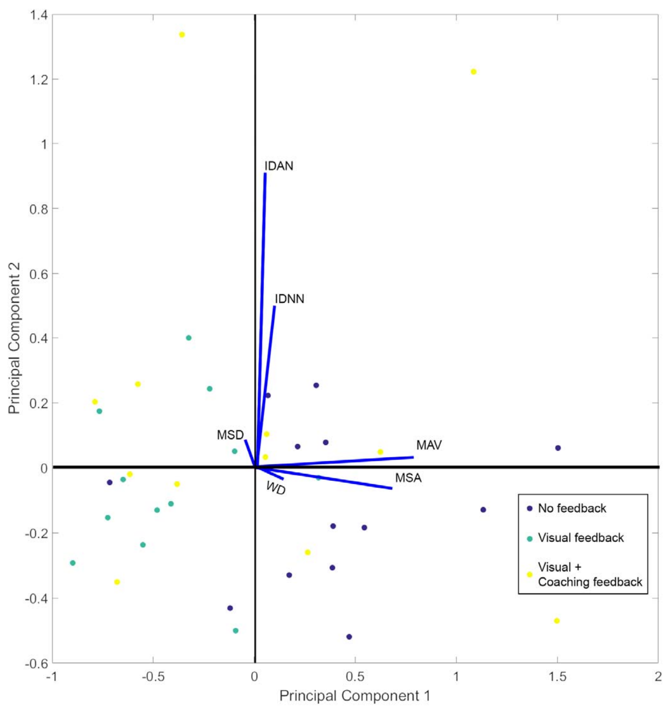

Fig. 7. Biplot of the first two principal components and loading vectors of the PCA. Each point represents the score of the first two principal components for one participant in a given group (NF/VF/VCF). The lines represent the first two principal component loading vectors. Abbreviations: IDNN (Interclass Distance Nearest Neighbor), IDAN (Interclass Distance All Neighbors), WD (Within-class Distance), MSA (Mean Semi-principal Axis), MAV (Mean Absolute Value), MSD (Most Separable Dimension).

C. Principal Component Analysis

Three PCs were analyzed explaining 61%, 23% and 8% of the variance, respectively. The composition of the first two PCs can be seen in Figure 7. The composition of the third PC is mainly explained by WD (0.93) and MSD (-0.28). The ANOVAs revealed that there was a significant effect of feedback group for PC1 ($F(2,34) = 5.863$; $p = .006$; $\eta^2_G = .13$). Post hoc tests revealed that participants in group NF had significantly larger values of PC1 than participants in group VF ($F(23)$; $p < .001$), but did not differ from VCF, see Figure 8.

No significant effects were found for PC2 and PC3.

D. Correlation Between Pattern Change and Training Outcome

The analysis revealed that there was no significant correlation between any primary and any secondary

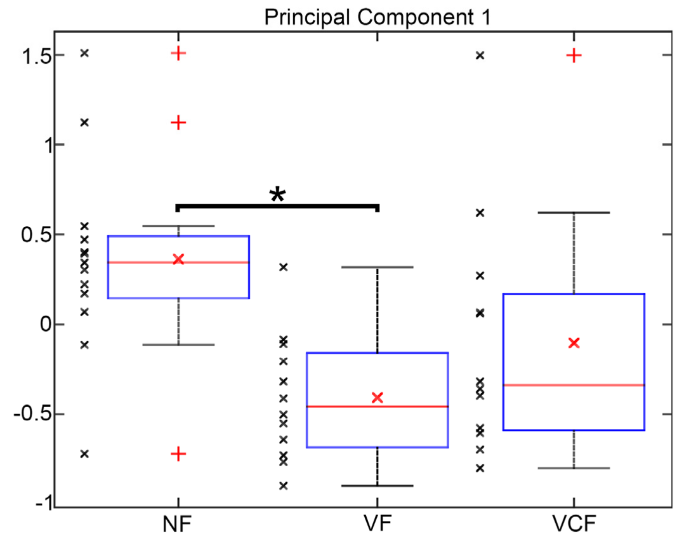

Fig. 8. Boxplot showing Principal Component 1 (PC1) from pre- to post-test. The red crosses denote the mean and the red line in each plot is the median. Outliers, marked with plus signs, are defined as points that lie outside 1.5 the interquartile range. Individual data points are marked with black crosses.

outcome variables. However, a correlation was found between online accuracy and PC1 ($\rho = .36$, $p = .028$). No significant correlation was found between the primary outcome variables and PC2 or 3.

IV. Discussion

To the best of our knowledge, this is the first study that compared different levels of feedback during user training in PR control over several days to determine the effect of feedback on control performance and EMG pattern quality. The results showed that user training improved PR control performance, independent of the level of feedback received during training. Training led to significant increases in performance (Figure 4a and 4b) that exceeds the ∼10% error differences between different classifiers [25], [26]. Training thus seems to have a larger impact on PR control than the choice of the classifier. For improving PR control, it might therefore be very beneficial to further investigate user training and its effect on classifier performance. In addition, we found that providing no external feedback during training led to higher amplitudes and more variability of the EMG patterns than training with feedback. Furthermore, a correlation was found between online accuracy and a principal component (PC1), primarily reflecting the amplitude and variance of the EMG patterns, derived from the feature space.

A. Training PR Control Is Possible Without Coaching

Our results demonstrated that user training for PR control led to significant improvements in performance even without visual feedback on the screen or individualized coaching. It might thus be possible to develop effective training schemes for the home environment, where therapist coaching is difficult to organize, in which higher training exposure is easily possible. This is important from a clinical standpoint, since insufficient training has been linked to high prosthesis abandonment rates [27], [28].

Our results are different from what we hypothesized and what would be expected from literature, where the study of Powell *et al.* [6] trains with high-feedback VCF and shows larger improvement after training than other studies training with less feedback [5], [12]. However, participants in the study by Powell et al. had an upper-limb defect, in contrast to the other studies [5], [12] in which able-bodied persons participated. Consequently, the persons with limb defects had the lowest pre-test scores and thus more room for improvement. However, a more comprehensive understanding of how differences in training effects between able-bodied and those with an upper-limb defect are mediated by feedback levels requires a direct comparison of those groups in an experiment. Such a comparison is currently not available in the literature and is something we aim to address in a future study.

B. Feedback Affects the Quality of EMG Patterns

We hypothesized that higher levels of feedback would lead to larger changes in the quality of the EMG, but it appears that this relation is not straightforward. The results show that training with no feedback increases EMG amplitudes (MAV) and pattern variability (MSA) compared to higher levels of feedback. Our results differ from those of He *et al.* [12] as we found no significant changes in pattern consistency (WD). This might be the result of the different pattern consistency metrics used or the longer training duration in the study by He et al. (10 days vs. 5 days in our study). Unlike the study of He et al., we additionally found a significant change in pattern separability. This effect that was nearly significant in the study of Bunderson and Kuiken [5], who reported that experienced participants had significantly higher pattern separability compared to trained novices. Our study confirms that training leads to higher pattern separability. It appears that different levels of feedback have different effects on the quality of the EMG patterns.

C. Correlation Between Outcome Metrics

We did not find a correlation between the primary and secondary outcome metrics. However, a compound metric derived from the secondary outcome metrics, namely principal component 1, was weakly correlated with online accuracy. While a compound metric, such as principal component 1, might be used to improve training programs, it is unlikely that the PCA space can be made generalizable enough to facilitate training for all PR control trainees.

Without a metric that correlates with performance it is difficult to facilitate training beyond trial and error. This is highly problematic since it is a requirement for designing an effective training scheme for PR control [5]. We also find it problematic for the widespread adoption of PR control that a strong relationship has not yet been established. This might also explain why all studies we know of comparing PR control with conventional direct control have been unable to clearly show that PR is superior to direct control [29]–[32] with the exception of one study performed on participants who had undergone targeted muscle re-innervation surgery [33]. Additional studies on user training for PR control should be conducted using different feedback metrics to find a correlation that can be used to facilitate training beyond trial and error.

The absence of a correlation between the primary and secondary outcome metrics does not mean that the primary and secondary outcome metrics are not related. Separability (Interclass Distance Nearest Neighbor, IDNN) and performance both increased significantly, but did not correlate with one another. This might indicate that certain metrics are not related to better performance beyond a specific value. For example, if patterns do not overlap, better separability may not lead to better performance. If a tool could be developed that can identify when EMG patterns are separable enough it would mean that the trainee would know when to shift focus from separating patterns to improving other aspects such as repeatability.

D. Participants Improve Using the Feedback That Is Available

It appears that all participants increased the separability of their EMG patterns (IDNN) from pre- to post-test using the feedback that was available to them. Participants in group NF appear to have done so by increasing the amplitude of their EMG. A suggestion that might explain why participants in the NF group generated EMG patterns with higher amplitudes than those in the other groups stems from the literature on motor learning [34]–[37]. It has been reported that having an internal focus (i.e. focus on one's movements, or one's muscle activity) as opposed to an external focus (i.e. focus on the results of one's movements) can lead to higher EMG amplitudes when learning motor skills. Participants in the NF group could only have used an internal focus since they did not have any external feedback (i.e. they could not see the result of their movements) and had to rely on the intrinsic feedback given by their muscles. The increase in EMG amplitude when training using an internal focus has been shown in activities ranging from bicep curls to dart throwing [34]–[37]. Interestingly, the presumed side effect of no feedback, which is using an internal focus, is the knowledge that increasing EMG amplitudes could be beneficial when learning PR based prosthesis control. While this was presumably the effect of participants in group NF having an internal focus, it could also be achieved by strength training. Strength training might thus be an overlooked path to improve PR control performance. While our results in this regard are inconclusive, we suggest that future research considers muscle strength, or a derivative thereof such as stump diameter, when assessing user PR control skill. If muscle strength does lead to better PR control, it could be a valuable tool in clinical rehabilitation.

Although the EMG patterns of group VF and VCF did not increase in amplitude as group NF, some trends can still be observed. EMG patterns of group VF were the only ones that did not trend towards an increase in amplitude following training (see Figure 6) and as a result the variability (MSA) of the EMG patterns was also low. This can be interpreted as that participants in group VF made their patterns less variable and became proficient in reproducing them. EMG patterns of group VCF varied a lot between participants considering the spread in Figure 7. This might be interpreted as that participants in the VCF group used different aspects of the feedback to improve. These findings show that feedback can be successfully given at different levels facilitating different kinds of training. A question for future research is if different levels of feedback can be combined to improve training. As an example, training could start with coaching in the clinic to introduce the concepts of PR control and give initial feedback on how to adapt movements to generate distinct patterns. When the user is at home, training could continue with visual feedback, to decrease pattern variability, and with no feedback to increase EMG amplitudes to make patterns more distinct. In this way, the specifics of each type of feedback are optimally exploited in rehabilitation practice. A first step towards such a training scheme would be to validate the results of this work in a population with an upper-limb defect. If these effects have been established, other aspects of training can be further scrutinized in future research, such as optimal intensity of training and variability in training.

E. Limitations

Our study has some limitations, most prominently that participants were able-bodied and the results cannot be directly applied to those with an upper-limb defect. In research on PR algorithms similar performance differences have been found in participants with or without an upper limb defect [38] (e.g. the algorithm giving the best performance for able-bodied individuals also gave the best performance for those with an upper limb defect), but this has not been established for user training. Another limitation is that the outcome task, the Motion Test, does not measure prosthetic use, but only controllability of the classifier output. Lastly, the test and training platform were the same which means participants might have improved during testing. We believe this did not have an effect on the results, as all participants performed the same tests and all training was conducted between the two tests.

V. Conclusion

The present study demonstrated that user training is a highly effective way to improve PR control, independent of the level of the feedback given during training. Training led to higher EMG pattern separability independent of feedback and to higher EMG amplitude when training with no feedback. However, we found no correlation between qualities of the EMG and performance, indicating that the mechanisms behind improvement in user performance are not straightforward. Future research should investigate the effect of training on individuals with upper limb defects with a focus on functional prosthetic performance. Such research could be important to develop effective training schemes that do not rely on trial and error. Furthermore, a metric which correlates the quality of EMG patterns with user performance would be an important step to improve training.

Acknowledgment

The authors would like to thank Alain Caspers, Erik Jan Nab, Joris Beintema and Teun van Otten for assisting with the experimental sessions.

References

[1] D. Farina and S. Amsüss, "Reflections on the present and future of upper limb prostheses," *Expert Rev. Med. Devices*, vol. 13, no. 4, pp. 321–324, 2016.
[2] J. L. Betthauser *et al.*, "Limb position tolerant pattern recognition for myoelectric prosthesis control with adaptive sparse representations from extreme learning," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 4, pp. 770–778, Apr. 2018.
[3] N. Nilsson, B. Håkansson, and M. Ortiz-Catalan, "Classification complexity in myoelectric pattern recognition," *J. Neuroeng. Rehabil.*, vol. 14, no. 1, p. 68, Jul. 2017.
[4] L. J. Hargrove, B. A. Lock, and A. M. Simon, "Pattern recognition control outperforms conventional myoelectric control in upper limb patients with targeted muscle reinnervation," in *Proc. 35th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2013, pp. 1599–1602.
[5] N. E. Bunderson and T. A. Kuiken, "Quantification of feature space changes with experience during electromyogram pattern recognition control," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 20, no. 3, pp. 239–246, May 2012.
[6] M. Powell, R. Kaliki, and N. Thakor, "User training for pattern recognition-based myoelectric prostheses: Improving phantom limb movement consistency and distinguishability," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 3, pp. 522–532, May 2014.
[7] J. Maier, A. Naber, and M. Ortiz-Catalan, "Improved prosthetic control based on myoelectric pattern recognition via wavelet-based de-noising," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 2, pp. 506–514, Feb. 2018.
[8] C. Castellini *et al.*, "Proceedings of the first workshop on peripheral machine interfaces: Going beyond traditional surface electromyography," *Frontiers Neurorobot.*, vol. 8, p. 22, Aug. 2014.
[9] R. E. Johnson, K. P. Kording, L. J. Hargrove, and J. W. Sensinger, "EMG versus torque control of human–Machine systems: Equalizing control signal variability does not equalize error or uncertainty," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 6, pp. 660–667, Jun. 2017.
[10] E. Martinez-Valdes, F. Negro, D. Falla, A. M. de Nunzio, and D. Farina, "Surface electromyographic amplitude does not identify differences in neural drive to synergistic muscles," *J. Appl. Physiol.*, vol. 124, no. 4, pp. 1071–1079, 2018.
[11] A. M. Simon, B. A. Lock, and K. A. Stubblefield, "Patient training for functional use of pattern recognition–Controlled prostheses," *J. Prosthetics Orthotics*, vol. 24, no. 2, pp. 56–64, 2012.
[12] J. He, D. Zhang, N. Jiang, X. Sheng, D. Farina, and X. Zhu, "User adaptation in long-term, open-loop myoelectric training: Implications for EMG pattern recognition in prosthesis control," *J. Neural Eng.*, vol. 12, no. 4, p. 046005, 2015.
[13] M. S. Cohen. (2008). *Brain Mapping, Handedness Questionnaire' Handedness Questionnaire*. Accessed: Jun. 3, 2018. [Online]. Available: http://www.brainmapping.org/shared/Edinburgh.php
[14] R. C. Oldfield, "The assessment and analysis of handedness: The Edinburgh inventory," *Neuropsychologia*, vol. 9, no. 1, pp. 97–113, 1971.
[15] A. L. Fougner, Ø. Stavdahl, and P. J. Kyberd, "System training and assessment in simultaneous proportional myoelectric prosthesis control," *J. Neuroeng. Rehabil.*, vol. 11, no. 1, p. 75, 2014.
[16] L. H. Smith, T. A. Kuiken, and L. J. Hargrove, "Myoelectric control system and task-specific characteristics affect voluntary use of simultaneous control," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 1, pp. 109–116, Jan. 2016.
[17] S. Amsuess *et al.*, "Context-dependent upper limb prosthesis control for natural and robust use," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 7, pp. 744–753, Jul. 2016.
[18] T. A. Kuiken *et al.*, "Targeted muscle reinnervation for real-time myoelectric control of multifunction artificial arms," *J. Amer. Med. Assoc.*, vol. 301, pp. 619–628, Feb. 2009.
[19] M. A. Powell and N. V. Thakor, "A training strategy for learning pattern recognition control for myoelectric prostheses," *J. Prosthetics Orthotics*, vol. 25, no. 1, pp. 30–41, 2013.
[20] A. D. Roche *et al.*, "A structured rehabilitation protocol for improved multifunctional prosthetic control: A case study," *J. Vis. Exp.*, vol. 105, Nov. 2015, Art. no. e52968.
[21] N. Nilsson and M. Ortiz-Catalan, "Estimates of Classification Complexity for Myoelectric Pattern Recognition," in *Proc. 23rd Int. Conf. Pattern Recognit. (ICPR)*, Dec. 2016, pp. 2682–2687.
[22] P. Roduit, M. Alcherio, and J. Jacques, "Behavioral analysis of mobile robot trajectories using a point distribution model," in *From Animals to Animats*. Berlin, Germany: Springer-Verlag, 2006, pp. 819–830.
[23] R. Bakeman, "Recommended effect size statistics for repeated measures designs," *Behav. Res. Methods*, vol. 37, no. 3, pp. 379–384, 2005.

[24] J. E. Jackson, *A User's Guide to Principal Components*. Chichester, U.K.: Wiley, 1991.

[25] L. J. Hargrove, K. Englehart, and B. Hudgins, "A comparison of surface and intramuscular myoelectric signal classification," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 5, pp. 847–853, May 2007.

[26] M. Ortiz-Catalan, R. Brånemark, and B. Håkansson, "BioPatRec: A modular research platform for the control of artificial limbs based on pattern recognition algorithms," *Source Code Biol. Med.*, vol. 8, no. 1, p. 11, 2013.

[27] E. Biddiss and T. Chau, "Upper-limb prosthetics: Critical factors in device abandonment," *Amer. J. Phys. Med. Rehabil.*, vol. 86, no. 12, pp. 977–987, 2007.

[28] M. L. McCarthy, T. Ewashko, and E. J. MacKenzie, "Determinants of use of outpatient rehabilitation services following upper extremity injury," *J. Hand Therapy*, vol. 11, no. 1, pp. 32–38, 1998.

[29] L. Resnik, H. H. Huang, A. Winslow, D. L. Crouch, F. Zhang, and N. Wolk, "Evaluation of EMG pattern recognition for upper limb prosthesis control: A case study in comparison with direct myoelectric control," *J. Neuroeng. Rehabil.*, vol. 15, no. 1, p. 23, 2018.

[30] E. Mastinu, J. Ahlberg, E. Lendaro, L. Hermansson, B. Håkansson, and M. Ortiz-Catalan, "An alternative myoelectric pattern recognition approach for the control of hand prostheses: A case study of use in daily life by a dysmelia subject," *IEEE J. Transl. Eng. Health Med.*, vol. 6, 2018, Art. no. 2600112.

[31] I. Vujaklija, S. Amsuess, A. D. Roche, D. Farina, and O. C. Aszmann, "Clinical evaluation of a socket-ready naturally controlled multichannel upper limb prosthetic system," in *Biosystems Biorobotics*, vol. 16. Cham, Switzerland: Springer, 2017, pp. 3–7.

[32] T. A. Kuiken, L. A. Miller, K. Turner, and L. J. Hargrove, "A comparison of pattern recognition control and direct control of a multiple degree-of-freedom transradial prosthesis," *IEEE J. Transl. Eng. Health Med.*, vol. 4, 2016, Art. no. 2100508.

[33] L. J. Hargrove, L. A. Miller, K. Turner, and T. A. Kuiken, "Myoelectric pattern recognition outperforms direct control for transhumeral amputees with targeted muscle reinnervation: A randomized clinical trial," *Sci. Rep.*, vol. 7, no. 1, Oct. 2017, Art. no. 013840.

[34] J. Vance, G. Wulf, T. Töllner, N. McNevin, and J. Mercer, "EMG activity as a function of the performer's focus of attention," *J. Motor Behav.*, vol. 36, no. 4, pp. 450–459, 2004.

[35] K. R. Lohse, D. E. Sherwood, and A. F. Healy, "How changing the focus of attention affects performance, kinematics, and electromyography in dart throwing," *Hum. Movement Sci.*, vol. 29, no. 4, pp. 542–555, 2010.

[36] D. C. Marchant, M. Greig, J. Bullough, and D. Hitchen, "Instructions to adopt an external focus enhance muscular endurance," *Res. Quart. Exerc. Sport*, vol. 82, no. 3, pp. 466–473, 2011.

[37] G. Wulf and R. Lewthwaite, "Effortless motor learning? An external focus of attention enhances movement effectiveness and efficiency," in *Effortless Attention: A New Perspective in Attention and Action*. Cambridge, MA, USA: MIT Press, 2010, pp. 75–101.

[38] E. Scheme and K. Englehart, "Electromyogram pattern recognition for control of powered upper-limb prostheses: State of the art and challenges for clinical use," *J. Rehabil. Res. Develop.*, vol. 48, no. 6, pp. 643–660, 2011.