Received 30 May 2017; revised 11 September 2017; accepted 22 October 2017. Date of publication 29 November 2017; date of current version 12 December 2017.

Digital Object Identifier 10.1109/JTEHM.2017.2776925

Evaluation of Computer-Based Target Achievement Tests for Myoelectric Control

JACOB GUSMAN[1], ENZO MASTINU[2], (Student Member, IEEE), AND MAX ORTIZ-CATALÁN[2,3], (Member, IEEE)

[1]Center for Biomedical Engineering, Brown University, Providence, RI 02912, USA
[2]Department of Electrical Engineering, Chalmers University of Technology, 412 96 Gothenburg, Sweden
[3]Integrum AB, 431 37 Mölndal, Sweden

CORRESPONDING AUTHOR: M. ORTIZ-CATALÁN (maxo@chalmers.se)

This work was supported in part by the Whitaker International Program, in part by the Swedish Research Council (Vetenskapsrdet) under Grant 621-2013-5749, in part by VINNOVA under Grant 2016-02290, and in part by the Promobilia Foundation.

ABSTRACT Real-time evaluation of novel prosthetic control schemes is critical for translational research on artificial limbs. Recently, two computer-based, real-time evaluation tools, the target achievement control (TAC) test and the Fitts' law test (FLT), have been proposed to assess real-time controllability. Whereas TAC tests provides an anthropomorphic visual representation of the limb at the cost of confusing visual feedback, FLT clarifies the current and target locations by simplified non-anthropomorphic representations. Here, we investigated these two approaches and quantified differences in common performance metrics that can result from the chosen method of visual feedback. Ten able-bodied and one amputee subject performed target achievement tasks corresponding to the FLT and TAC test with equivalent indices of difficulty. Able-bodied subjects exhibited significantly ($p < 0.05$) better completion rate, path efficiency, and overshoot when performing the FLT, although no significant difference was seen in throughput performance. The amputee subject showed significantly better performance in overshoot at the FLT, but showed no significant difference in completion rate, path efficiency, and throughput. Results from the FLT showed a strong linear relationship between the movement time and the index of difficulty ($R^2 = 0.96$), whereas TAC test results showed no apparent linear relationship ($R^2 = 0.19$). These results suggest that in relatively similar conditions, the confusing location of virtual limb representation used in the TAC test contributed to poorer performance. Establishing an understanding of the biases of various evaluation protocols is critical to the translation of research into clinical practice.

INDEX TERMS Electromyography (EMG), fitts' law, myoelectric control, target achievement control (TAC), user interfaces.

I. INTRODUCTION

AMONG the various efforts to restore functional capacity to limb amputees, myoelectric pattern recognition (MPR) has emerged as a clinically viable technique to address the challenge of intuitive control of a limb prosthesis [1]. Unlike the experience of operating external, man-made devices, the experience of natural bodily movement is practiced, developed, and familiarized since infancy. It is nuanced and intuitive, and thus, any meaningful replacement is expected to satisfy a premium standard of quality that we do not expect of non-biomimetic systems. As such, numerous variations on MPR, from electrode positioning [2] to classification schemes [3]–[6], have been proposed as improvements toward the goal of better prosthetic control. However, in this relatively young field, few generally accepted tools exist to quantify these improvements.

In many MPR studies, *classification accuracy* is used as the primary metric for determining system performance [1]. A large variety of classification techniques have been proposed [4], [7], including the most commonly used supervised classifiers of Linear Discriminate Analysis (LDA) and Multi Layer Perceptron (MLP). Pattern recognition algorithms are typically assessed offline by measuring the accuracy of class (movement/posture) estimates that the classifier makes when presented with new, unseen data.

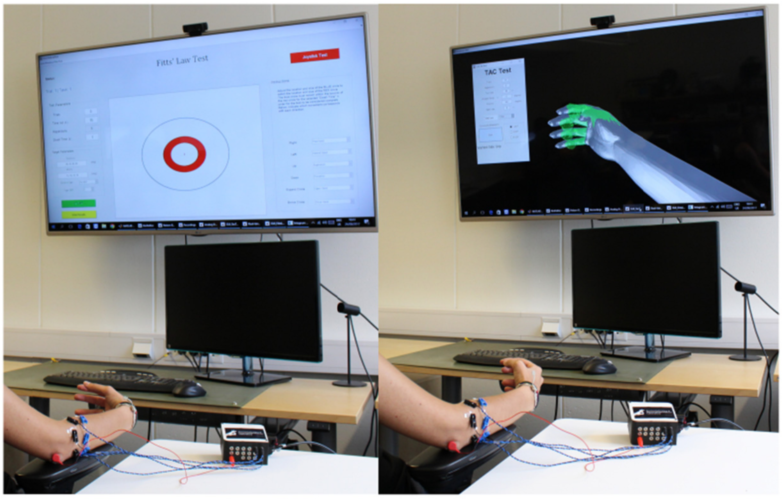

FIGURE 1. **The experimental setup and graphical user interfaces (GUI) used for visual feedback in the FLT (left) and TAC test (right).**

Although classification accuracy remains an important initial assessment, a substantial argument has been made disputing the correlation of offline assessment to online performance [8]–[11]. Hargrove *et al.* [12] showed that the inclusion of transient contractions in training data, which lowered classification accuracy, improved subject performance in a real-time virtual clothespin task.

In an attempt to more faithfully assess real-life performance, researchers have adapted existing tests from the rehabilitation field to evaluate prosthesis control in more clinically relevant scenarios. The box and blocks (B&B) test [13], for example, is among the simplest of these tests, requiring subjects to individually transfer square blocks over a barrier in a set amount of time. For a more comprehensive assessment, researchers have employed the Southampton Hand Assessment Protocol (SHAP) [14], a test consisting of 26 distinct tasks and 6 grips. SHAP, however, is often seen as too lengthy and tiring for many patients [11], and, similar to the B&B test, relies on a performance metric of completion time that fails to characterize the full movement path. A clinical tool designed specifically for the evaluation of myoelectrically-controlled prostheses is the Assessment for Capacity of Myoelectric Control (ACMC) [15], [16], the use of which has been limited potentially due to its fairly large subjective component [11]. Furthermore, whereas these clinical tests succeed in incorporating influential environmental factors such as socket instability and sweating, the additional complexity and reliance on the presence of certified examiners and materials could make such protocols burdensome for many research groups.

Recently, several alternatives to the above offline and clinical approaches have emerged in the form of computer-based virtual assessment tools. One of the simplest of these is called the Motion Test. The Motion Test was designed by Kuiken *et al.* [17] to evaluate the myoelectric control capacity

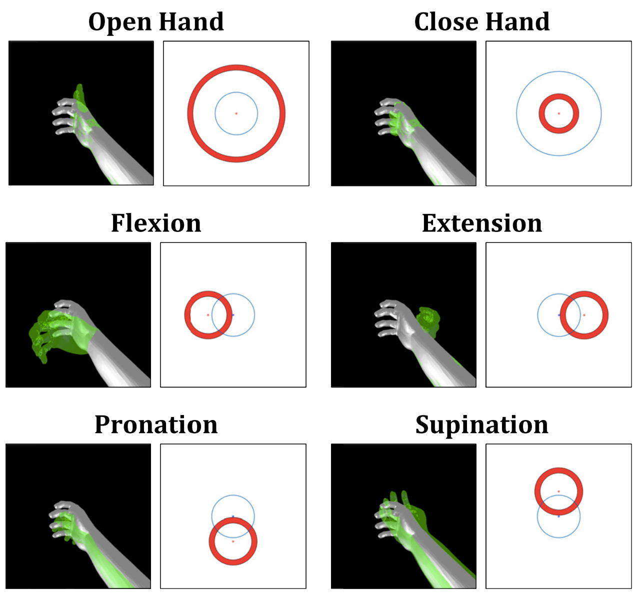

FIGURE 2. **Examples of the visual feedback provided to subjects performing the TAC test (left) and the FLT (right) for each of the six movements. Distance (D) = 60° and Width (W) = 13°.**

of patients who underwent a targeted muscle reinnervation (TMR) procedure. The test is initiated by presenting to the patient a visual prompt requesting that they attempt one of a set of arm motions that they have previously trained with a classifier. Once the patient has selected the requested movement, they must maintain that contraction until the classifier has made a predetermined number of correct predictions. This can also be visually represented by a virtual arm on a screen traveling through the full range of the movement.

Although a significant step in the direction of a more realistic testing scenario, the Motion Test, according to Simon *et al.* [18], is ''oversimplified'', and an inadequate test for evaluating control schemes which allow misclassification or speed modulation. Simon *et al.* proposed a more comprehensive test called the Target Achievement Control (TAC) test. In the TAC test (Fig. 1, right inset), the user has multidimensional control of a virtual prosthesis and is prompted to move the virtual prosthesis toward a target posture, where it must remain for a predetermined *dwell time* (i.e. 1 s). While performing the task, the user's progress can be impeded by unintended misclassifications and target overshoots – similar to how a myoelectric prosthesis might perform in real life. The test can be customized to restrict the dimensional freedom of the virtual prosthesis and/or alter the complexity of the task by changing the location and/or width of the target. Among the performance metrics assessed by Simon *et al.* [18] were completion rate, completion time, and path efficiency.

Upon examination, Scheme and Englehart [19] made the correct observation that the TAC test resembled a better known test used in the evaluation of other Human Computer Interfaces (HCI) called the Fitts' Law Test (FLT). Originally demonstrated by Fitts in 1954 [20], the Fitts' Law Test has become an international standard (ISO9341-9) for the validation of practically any type of human-computer interface (HCI) including mice, joysticks, touchpads, and even eye-trackers [21]. However, only recently has EMG been shown to be a suitable control source for Fitts' Law testing [22]–[24]. At its core, Fitts' Law asserts that there exists a trade-off between speed and accuracy in target acquisition tasks that is defined by

$$MT = a + b * ID \tag{1}$$

where MT is the time (in seconds) required to acquire the target, a and b are regression coefficients, and ID is the target's *index of difficulty* (in bits), which is defined as

$$ID = log_2\left(\frac{D}{W} + 1\right) \tag{2}$$

where D and W are the target distance and width, respectively. Fitts proposed a metric called *throughput (TP)* that could quantitatively describe the performance of a given control system:

$$TP = \left(\frac{ID}{MT}\right) \tag{3}$$

Thus, Scheme et al. devised a pseudo 3-D Fitts' Law style task that accepts myoelectric input to perform target acquisition tasks on a screen (Fig. 1, left inset). The protocol is similar to that of the TAC test, but instead of a virtual arm, tasks are performed using a circular cursor capable of three-degrees of freedom. This simplification was seen as a benefit to Scheme at al., citing the ''limitations in performance, visualization, and immersion'' of virtual environments that attempt to more realistically represent the movement of a prosthetic device.

The FLT and the TAC test aim to satisfy an important need in the development of novel prosthetic control techniques. However, no existing studies have attempted to compare these evaluation tools directly. In this study, the FLT and TAC test were performed under equivalent conditions in order to investigate and quantify the difference in their outcome metrics. The data collected in this work, as well as the implementation of both real-time tests, had been made freely available as open source as part of BioPatRec (release FEM – to be released with the publication of this article) [7].

II. METHODS

A. SUBJECTS

Ten able-bodied subjects and one amputee participated in this study. Of the 10 able-bodied subjects, 5 were female, 2 were left-handed, and 7 had no experience using a myoelectric pattern recognition system. All subjects ranged in age from 20 to 44. The amputee subject had a left transradial amputation, and regularly uses a myoelectric prosthesis. All the experiments were approved by the Swedish Regional Ethics Committee in Gothenburg (626-10, T688-12).

B. RECORDING PROCEDURE

Surface EMG signals throughout the experiment were acquired using a custom acquisition device based on the RHA2216 analog-front-end (Intan Technologies, USA) [25] The signal was sampled at 2 kHz frequency and subsequently filtered using a 4^{th} order Butterworth filter with a 20-800 Hz passband. Four Ag-AgCl bipolar electrode pairs were evenly distributed around the most proximal third of the subject's dominant forearm, and one reference electrode was placed at the elbow. Signals were preliminarily inspected and electrode positions were adjusted slightly if necessary to achieve acceptable signal quality.

Recording sessions involved 7 movements: hand open/close, wrist flexion/extension, pro/supination, and rest (no movement). Subjects were asked to execute and maintain each movement at ~70% full strength for 3 seconds and rest for 3 seconds, iterated 3 times.

C. SIGNAL PROCESSING AND PATTERN RECOGNITION

All procedures for signal processing and pattern recognition were enabled by the open source platform BioPatRec [7]. Fifteen percent of the signal from each contraction time was discarded from the beginning and the end in order to remove idle periods while partially preserving transient portions of the contraction signal. The resulting EMG data was evaluated using overlapping time windows of 200ms in 50ms increments. From each time window, the four most commonly used time-domain features were extracted [5]: mean absolute value, zero crossings, slope sign changes, and waveform length.

The linear discriminant analysis (LDA) pattern recognition algorithm in the one-vs-one (OVO) topology was used to classify the feature values, a method that has been shown previously to be an effective strategy for the real-time control of individual movements [19], [26], [27].

D. REAL-TIME TESTS

The TAC test, which was originally proposed by Simon *et al.* [18], was previously incorporated into BioPatRec [27], an open source, MATLAB-based (The MathWorks, Inc., Natick, MA), research platform for advanced myoelectric control [7]. In this study, the TAC test implementation in BioPatRec was adapted to allow for the random presentation of targets of different distances (°) and allowances (°, equivalent to one half the target width) in the same trial (Fig. 1 right inset).

The Fitts' Law Test utilized in this study was modeled after the pseudo-3D task presented by Scheme & Englehart [19] and implemented from scratch in BioPatRec (Fig. 1 left inset). Instead of a virtual limb like in the TAC test, the users controlled a circular cursor in three dimensions aiming to match the location and radius of the target.

In this study, hand flexion and extension were mapped to the leftward and rightward movements of the cursor (rightward and leftward for left-handed subjects), supination and pronation were mapped to the upward and downward movement of the cursor, and open hand and close hand were mapped to the increase and decrease of the cursor radius (Fig. 2).

Similar to the TAC test, subjects were asked to move the cursor, which originated at the center of the screen with a medium radius, to the location and/or size of the target cursor, and remain within the *width* of the target for the predetermined *dwell time*. All attempts were made to make sure that the fundamental parameters of the FLT task including target distance, target width, speed, and dwell time, were represented identically to the parameters of the TAC test. The implementations of the TAC test and FLT were designed to use the same distance units, such that if, for example, a target in the TAC test requires 100 consecutive "wrist flexion" classifications for completion, that same target in the FLT also requires 100 consecutive "wrist flexion" classifications for completion. In this way, the only difference between the two tests was the visual representation of user movement on the screen.

E. EXPERIMENTAL PROTOCOL

Subjects were asked to perform trials of both FLT and TAC test. In order to reduce bias due to the learning effect, half of the subjects were randomly selected to begin with the FLT and half to begin with the TAC test. Furthermore, subjects were given a training trial before the first trial of each test in order to familiarize themselves with the task, and subjects alternated between trials of the FLT and the TAC test in order to control for fatigue. In this study, each *trial* consisted of a set of 24 target acquisition tasks.

In both the TAC test and the FLT, the hand/cursor was free to move in all 3 degrees of freedom (DoF), but targets were limited to one DoF at varying IDs. Therefore, each target only required one type of movement to be reached. If an overshoot occurred, or if an unintended movement was produced, other types of movements were necessary to correct for the error.

TABLE 1. Sets of distances (D) and widths (W) and resulting indices of difficulty (ID).

Distance (D)	Width (W)	Index of Difficulty (ID)
30°	13°	1.73 bits
30°	8°	2.25 bits
60°	13°	2.49 bits
60°	8°	3.09 bits

It was suggested in [19] that Fitts' Law is best examined utilizing a variety of ID values under 4 bits. Accordingly, four sets of distances (D) and widths (W), which are shown in Table 1, were used during this study.

For both tests, the velocity ramp [5] postprocessing strategy was used with a ramp length of 10 and a down count of 2. The maximum allowed displacement was 2 degrees per prediction, which, when paired with the 50 ms time increment, results in a maximum speed of 40 degrees per second. Subjects were expected to complete each target acquisition task in less than the 15-second time out time or be marked incomplete. Dwell time was set to 1 second.

Subjects performed 3 trials of each test (excluding one training trial for each test). Each trial consisted of 24 randomly presented target acquisition tasks, representing all distinct combinations of the 6 target directions (open/close hand, hand flexion/extension, pro/supination) and 4 indices of difficulty (ID) (Table 1). Each subject, therefore, performed 72 target acquisition tasks per test.

F. PERFORMANCE METRICS

In addition to throughput, performance metrics of completion rate, efficiency, and overshoot, were calculated [19]. *Throughput* was defined in Equation 3, with possible IDs shown in Table 1. *Movement time* was defined as the time between the beginning of a task and the initial acquisition of the target (i.e. completion time – dwell time) or the time out time, in the case of incomplete tasks. *Completion rate* was defined as the percentage of tasks completed by the subject within the 15-second time out time. *Path efficiency* (%) was determined for each task and represents the ratio of the shortest distance to the target to the actual distance travelled. The *overshoot* metric refers to the number of times per task that a target was acquired but then lost before the dwell time was reached.

Alternative methods for calculating ID and MT have been proposed previously [17], [28]) and were implemented here in order to determine whether such variations on calculation procedure had a significant effect. Soukoreff & Mackenzie recommended using an *effective target width* (*We*) when constructing Fitts' Law models [28], which can be approximated

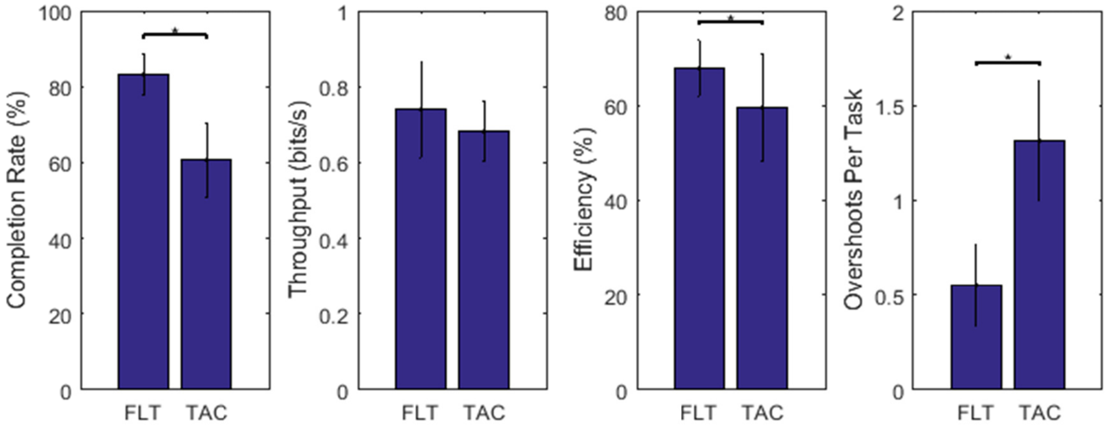

FIGURE 3. **Resulting performance metrics for FLT and TAC when tested under the same conditions in able-bodied subjects. Compared to FLT, TAC had a significantly lower completion rate (%) and efficiency (%), and a significantly higher occurrence of overshoots (per task). Results for throughput (bits/s) were not significantly different. Errorbars indicate standard deviations and * indicates a significant difference (p < 0.05).**

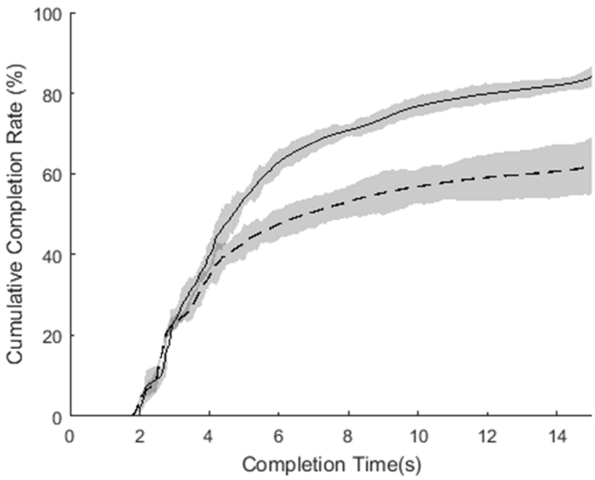

FIGURE 4. **Cumulative completion rate curves for FLT (solid line) and TAC (dashed line) tests in able-bodied subjects. Shaded regions represent ±1 standard error.**

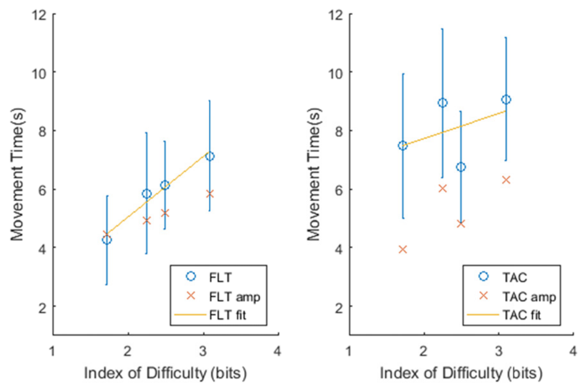

FIGURE 5. **The relationship between movement time (s) and index of difficulty (bits) for both the Fitts' Law Test (left) and the Target Achievement Control test (right) for ID = 1.73, 2.25, 2.49, and 3.09 bits. Blue circular points and error bars indicate the mean and standard deviation, respectively, of results from 10 able-bodied subjects. The solid yellow line is a regression of the ablebodied results. The red x's show the results from the amputee tested.**

using the error rate (Err) of each condition:

$$We = \begin{cases} W \times \frac{2.066}{z(1-Err/2)} & ifErr > 0.0049\% \\ W \times 0.5089 & otherwise. \end{cases} \tag{4}$$

where $z(x)$ is the inverse of the normal cumulative distribution function, and

$$IDe = log_2\left(\frac{D}{We}+1\right) \tag{5}$$

where IDe is the *effective index of difficulty*. This method is meant to incorporate the accuracy variations that may exist between different subjects and conditions [28].

Furthermore, another alternative for calculation of ID and MT was explored wherein the median values for this data were used instead of the mean, as Kuiken*et al.* did when evaluating the results of their Motion Test [17].

G. STATISTICAL ANALYSIS

Analysis of all collected data was conducted using MATLAB. A paired t-test was used to assess the statistical difference in the aforementioned performance measures between the two tests. The overall significance of linear regression models was determined using an F-test. A two-sample t-test was used to assess differences in performance between novices who began with the FLT and novices who began with the TAC test.

III. RESULTS

Fig. 3 summarizes the performance (mean ± standard deviation) of the 10 able-bodied subjects. The average completion rate (%) for trials of the FLT (83.19 ± 5.54) was significantly ($p < 0.001$) higher than that of the TAC test trials (60.69 ± 9.7). Throughput values of the FLT trials (0.52 ± 0.06) were slightly higher than the throughput of the TAC

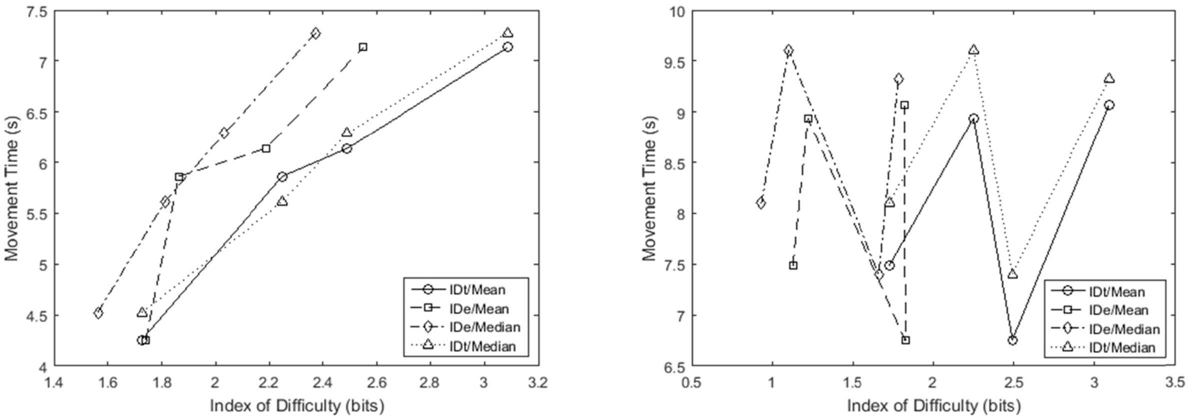

FIGURE 6. Different ways of plotting Index of Difficulty (ID) vs Movement Time (MT) for FLT (left) and TAC (right). The tested ID values (IDt) are shown in Table 1. Effective ID values (IDe) were estimated using Equation (5). Average MT's and ID's show up differently depending on whether the mean or median was used to calculate them. Linearity was assessed using the coefficient of determination, R^2 (FLT: $R^2_{IDt/Mean} = 0.960$, $R^2_{IDe/Mean} = 0.815$, $R^2_{IDe/Median} = 0.989$, $R^2_{IDt/Median} = 0.992$; TAC: $R^2_{IDt/Mean} = 0.187$, $R^2_{IDe/Mean} = 0.011$, $R^2_{IDe/Median} = 0.006$, $R^2_{IDt/Median} = 0.109$).

trials (0.46 ± 0.07), however this was not statistically significant ($p = 0.13$). The path efficiency (%) was significantly ($p = 0.02$) higher for the FLT trials (67.87 ± 5.87) than for the TAC test trials (59.47 ± 11.17). Results for overshoot indicate a more than double occurrence of overshoot in trials of the TAC test (1.32 ± 0.31) versus trials of FLT (0.55 ± 0.22, $p < 0.001$).

Fig. 4 displays the cumulative completion rate curves for the FLT and TAC tests in able-bodied subjects. By the timeout time of 15 seconds, 83.19 ± 5.54% of FLT tasks were completed and 60.69 ± 9.7% of TAC tasks were completed. The curves appear to "flatten out" well before the timeout time, suggesting that 15 seconds was a sufficient amount of time to assess completion rates.

Fig. 5 displays the relationship between movement time (MT) and index of difficulty (ID) for both the FLT and TAC test. Although both tests suggest a linear relationship between MT and ID, only that of the Fitts' Law Test can be said to be a strong relationship ($R^2 = 0.96$) that is statistically significant ($p = 0.02$). Results of the TAC test do not display a significant linear relationship between MT and ID ($R^2 = 0.19$, $p = 0.56$), indicating that Fitts' Law may not be an appropriate model for this type of test.

In Fig. 6, ID and MT are plotted using combinations of the alternative calculation methods described by Soukoreff and Mackenzie [28] and Kuiken *et al.* [17]. Despite the alternative calculations, the coefficients of determination, R^2, remained high for the FLT ($R^2_{FLT} > 0.8$) and low for the TAC test ($R^2_{TAC} < 0.2$), indicating that the results of Fig. 5 are not significantly affected by the use of these alternative calculation methods.

Fig. 7 and 8 display the path traces for all tests performed by a representative subject where D = 60 and W = 8 or 13. A straight line between the origin and a shaded target box would indicate a 100% efficiently performed task, whereas deviations from that line indicate errors. Lines that pass into and out of a target box indicate the occurrence of an overshoot.

During testing it was observed that novice users (i.e. subjects with no experience using MPR) who began the experiment learning the Fitts' Law Test (S4, S7, S9; trial pattern: 1. TAC 2. FLT 3. TAC 4. FLT 5. TAC 6. FLT) had an easier time learning the necessary contraction patterns, and thus they exhibited better performance in both the FLT and TAC test when compared with subjects who began the experiment learning the TAC test (S1, S3, S5, S8; trial pattern: 1. FLT 2. TAC 3. FLT 4. TAC 5. FLT 6. TAC). The results displayed in Fig. 9 suggest that the novices who began with the TAC test indeed had significantly lower completion rates when performing the TAC test than those who began with the FLT ($p_{TAC} = 0.044$). However, the test with which a subject began appeared to have no significant effect on the completion rate when performing the TAC test ($p_{FLT} = 0.725$).

A. AMPUTEE RESULTS

Table 2 presents the performance metrics produced from the transradial amputee subject's trials. Unlike the able-bodied participants, the amputee exhibited no difference in performance between the FLT and TAC test in the metrics of completion rate ($p = 0.78$), throughput ($p = 0.96$), and path efficiency ($p = 0.67$). Similar to the able-bodied subjects, the amputee made far more overshoots while doing the TAC test ($p < 0.001$).

The amputee subject, who was an experienced MPR user, performed better than most of the able-bodied subjects, and in 5 out of the 8 mean values shown in Table 2, the amputee performed significantly better than the other subjects ($p < 0.05$). Most notably, the subject performed significantly better in 3 out of 4 of the TAC metrics. However, when compared with the three other experienced MPR users (S2, S6, S10), only 2 out of the 8 mean values (FLT throughput and FLT efficiency) were significantly different.

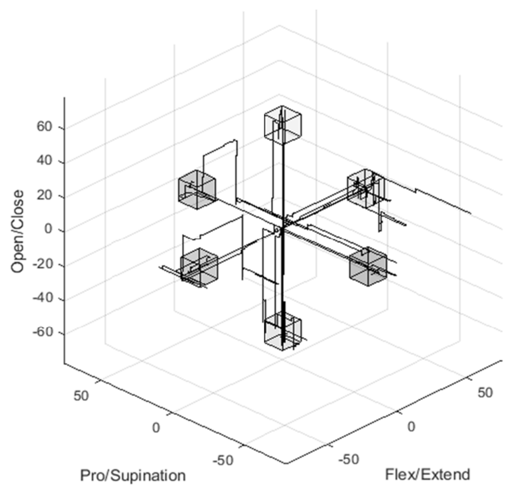

FIGURE 7. **Path traces from a representative subject using the Fitts' Law Test. Traces from tasks where D = 60° and W = [8°, 13°] are shown. Shaded boxes represent targets of width 13°. All axes units are in degrees (°) from the initial position.**

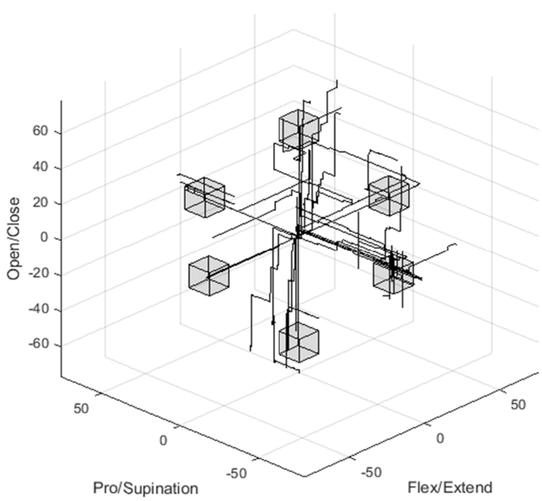

FIGURE 8. **Path traces from a representative subject using the TAC test. Traces from tasks where D = 60° and W = [8°, 13°] are shown. Shaded boxes represent targets of width 13°. All axes units are in degrees (°) from the initial position.**

IV. DISCUSSION

In prosthetics, computer-based target acquisition tests have proven to be convenient stand-ins for the more cumbersome real-life tests used in clinical settings. They also offer much more information on an algorithm's translational potential than offline measures such as classification accuracy.

Here we have examined two target acquisition tasks with computer-based implementations: the TAC test, which uses a virtual limb to simulate control of a prosthesis, and the FLT, a pseudo-3D Fitts' law style task where a circular cursor is manipulated on the screen. Despite having been tested under equivalent conditions, such as speed, target type, and dwell time, the resulting performance metrics of completion rate, path efficiency, and overshoot showed statistically significant differences (Fig. 3, Table 2).

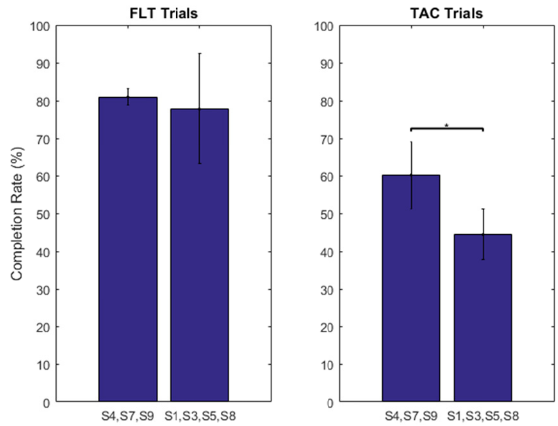

FIGURE 9. **Average completion rates across all trials for novice subjects who began the experiment with FLT (S4, S7, S9) and novice subjects who began the experiment with TAC (S1,S3,S5,S8). Errorbars indicate standard deviations and ∗ indicates a significant difference ($p < 0.05$).**

TABLE 2. **Results of FLT and TAC tests in amputee subject.**

	FLT	TAC	P-value
Completion Rate (%)	84.72 ± 36.23	86.11 ± 34.83*	p = 0.78
Throughput	0.62 ± 0.33*†	0.62 ± 0.35*	p = 0.96
Efficiency (%)	70.19 ± 34.38†	72.20 ± 32.40*	p = 0.67
Overshoot	0.25 ± 0.52*	1.31 ± 0.99	p < 0.001

*Two-sample *t*-test indicates significant difference between result of amputee subject and result of all able-bodied subjects (p<0.05)
†Two-sample *t*-test indicates significant difference between result of amputee subject and result of experienced, able-bodied subjects (p<0.05)

The most obvious difference found was the overshoot metric, where there were over twice as many overshoots in the TAC trials than in the FLT trials. This behavior can explain why the completion rate and path efficiency metrics were lower for the TAC test as well. Increased overshoots would require the subject to take a greater amount of time and move a greater total distance in order to complete the task. These results reflect the observations that participants voiced, saying that the target was difficult to locate in the TAC test, and once they found themselves close to the target, where the virtual target and controlled arms were overlapping, it was difficult to determine which movement was required to reach the target. This is shown visually by the path traces of the representative subject shown in Figs. 7 and 8, where the path traces of the TAC trials appear to diverge noticeably more from the ideal path to the target than do the traces of the FLT trials. Indeed, this confusion was a motivation for Scheme and Englehart [19] for devising a simplified abstraction of the task using a circular cursor.

The results for the FLT in Fig. 5 confirmed Scheme *et al.*'s assertion that their myoelectrically controlled pseudo-3D Fitts' law style task indeed follows the Fitts' Law model. However, similar analysis of the TAC test (Fig. 5, right inset) suggests that Fitts' Law testing may not be suitable for all forms of myoelectric control evaluation. This conclusion was supported by the indifference of the results to alternative calculation techniques for MT and ID (Fig. 6).

A notable difference between the visualizations offered by the FLT and TAC test is the visual representation of target width. Because of the reliance on using an arm-shaped target, the TAC test is unable to provide a clear representation of target width, whereas in the FLT target width is simply shown by the width of the target cursor. Although the target widths were in effect identical, the lack of visual feedback in the TAC test may have been what prompted the subjects to overshoot more frequently. This difference may also help to explain the TAC results of Fig. 5, where movement time appears to be dependent entirely on target width (i.e. MTs of targets with equal Ds but different Ws were significantly different, $p < 0.05$, but MTs of targets with different Ds but equal Ws were not statistically different, $p > 0.05$). The increase in difficulty associated with a smaller target (smaller W) is an assumption of the Fitts' Law model whose effect may be amplified if the change of target width is not properly represented in the visual feedback, as is the case in the TAC test. Moreover, this ability to show different widths represents an advantage to the FLT, but the usefulness of this feature can be debated, given that most ''targets'' in the real world are not colorful circles of various widths, but instead, physical objects that can require indistinct levels of precision.

This study emphasizes the importance of utilizing several performance metrics in myoelectric controls research. Although no significant difference was shown between the FLT and the TAC test in throughput, the two tests differed significantly when comparing completion rate, path efficiency, and overshoot. Therefore, multiple metrics should be used when describing control. Furthermore, as shown by the differences between the FLT and TAC test, each of these metrics may be biased depending on the nature of the test itself.

During the early trials of a testing session, it was observed that users new to myoelectric control (''novices'') had more difficulty learning to move the TAC arm than the FLT cursor, which was surprising given that the TAC test appears to display a more realistic, intuitive visualization of arm movement. However, during TAC trials, subjects would often attempt to simply match the posture of the requested movement rather than perform the correct series of contractions to ''move'' the arm into the correct position. For example, in a task that had a target at 30° flexion, a subject would be tempted to simply position their hand at 30° flexion, rather than fully flex their wrist and maintain the contraction until completion. Conversely, it appeared to be easier to teach the subject to perform the correct contractions when the 6 movements were matched to the 6 directions used to control the FLT cursor. Indeed, the results of Fig. 9 indicate that subjects who began the experiment by learning how to perform the FLT, performed significantly better on the TAC trials than those who began the experiment by learning the TAC test. This may be expected, since training initially with the FLT meant that a subject had more overall training experience before beginning the first TAC test. However, if this were the case, it would also be expected that subjects who began with the TAC test would perform better on the FLT trials than those who began with the FLT, yet no significant difference in completion rate of the FLT trials was observed between novices who began with the FLT and novices who began with the TAC test. This result suggests not only that the FLT is easier to learn for novices but also that experience learning the FLT may increase performance in more confusing target acquisition visualizations like what was used in the TAC test.

The fact that the FLT represents movement using a cursor, and the TAC test uses a virtual representation of the arm, suggests that the FLT can be more easily equated to a ''game''; one subject even remarked that the FLT was more ''fun because it looks like a game''. Indeed, much study has been devoted to determining the ability of myoelectrically controlled games to translate skills to real-world prosthesis use. Van Dilk *et al.* showed that previous training using an EMG-controlled game had no effect on performance in a prosthesis-simulator task [29]. Augmented reality training, on the other hand, has been shown to improve myoelectric prosthesis training [30]. Although both the TAC test and the FLT can be used as an indication of real-time controllability, further work needs to be done to understand to what degree each test can describe performance in the clinical setting.

Many limb movements fall out of the realm of target acquisition. Using tools like pencils or spoons or performing gestures like waving or even dancing is dynamic; success is determined not by start and end point, but by the entire path itself (Fig. 7 and 8). The FLT and TAC rank performance of control based on average times and distances. Differences in standard deviation are often not considered, and there is no difference between ten 1° errors and one 10° errors. Moreover, smoothness of movement and the control of acceleration and deceleration of a prosthesis, which can be important in sensitive situations such as carrying a plate of food or a boiling pot of water, have not been accounted for yet in these tests.

Finally, although it was found that results of the amputee tested in this study did not vary significantly from those of the other subjects, more amputee subjects must be included in future trials to conclusively state that evaluation of able-bodied subjects using FLT and TAC is translatable to the amputee population.

V. CONCLUSION

A direct comparison was performed between two common computer-based assessment tools for the evaluation of myoelectric control. Ten able-bodied subjects and one amputee were tested using a pseudo-3D Fitts' Law Test (FLT) and the Target Achievement Control (TAC) test under identical conditions. Subjects performed significantly worse in the TAC test with regard to completion rate ($p < 0.001$), path efficiency ($p = 0.02$), and overshoot ($p < 0.001$), but showed no significant difference in throughput ($p = 0.13$). Regression plots showed that the FLT adhered closely to the

Fitts' Law model, whereas the TAC test did not. The lack of target width visualization in the TAC test was identified as a contributor to its poorer performance.

Despite their inherent similarities, the TAC test and the FLT cannot be considered interchangeable. Our testing showed that the VR environment used in the TAC test is responsible for a significant increase in user error and reported confusion, suggesting that the FLT may be a more reliable assessment tool, especially when testing using multiple target widths. However, further investigation is required to determine which test has greater potential for translation to the clinical environment.

ACKNOWLEDGMENT

The authors would like to thank all study participants. J. Gusman was with the Biomechatronics and Neurorehabilitation Laboratory, Department of Electrical Engineering, Chalmers University, Gothenburg, Sweden.

REFERENCES

[1] 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. Dev.*, vol. 48, no. 6, pp. 643–660, 2011.

[2] 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.

[3] Y. Huang, K. B. Englehart, B. Hudgins, and A. D. C. Chan, "A Gaussian mixture model based classification scheme for myoelectric control of powered upper limb prostheses," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 11, pp. 1801–1811, Nov. 2005.

[4] E. J. Scheme, K. B. Englehart, and B. S. Hudgins, "Selective classification for improved robustness of myoelectric control under nonideal conditions," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 6, pp. 1698–1705, Jun. 2011.

[5] K. Englehart and B. Hudgins, "A robust, real-time control scheme for multifunction myoelectric control," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 7, pp. 848–854, Jul. 2003.

[6] M. A. Oskoei and H. Hu, "Support vector machine-based classification scheme for myoelectric control applied to upper limb," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 8, pp. 1956–1965, Aug. 2008.

[7] 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. 11, pp. 1–18, 2013.

[8] B. A. Lock, K. Englehart, and B. Hudgins, "Real-time myoelectric control in a virtual environment to relate usability vs. accuracy," in *Proc. Myo-Electric Controls Symp.*, 2005, pp. 122–127.

[9] N. Jiang, S. Dosen, K.-R. Muller, and D. Farina, "Myoelectric control of artificial limbs—Is there a need to change focus?" *IEEE Signal Process. Mag.*, vol. 29, no. 5, pp. 150–152, Sep. 2012.

[10] M. Ortiz-Catalan, F. Rouhani, R. Branemark, and B. Hakansson, "Offline accuracy: A potentially misleading metric in myoelectric pattern recognition for prosthetic control," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 1140–1143.

[11] I. Vujaklija *et al.*, "Translating research on myoelectric control into clinics-are the performance assessment methods adequate?" *Frontiers Neurorobot.*, vol. 11, no. 7, pp. 1–7, Feb. 2017.

[12] L. Hargrove, Y. Losier, B. Lock, K. Englehart, and B. Hudgins, "A real-time pattern recognition based myoelectric control usability study implemented in a virtual environment.," in *Proc. 29th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2007, pp. 4842–4845.

[13] V. Mathiowetz, G. Volland, N. Kashman, and K. Weber, "Adult norms for the box and block test of manual dexterity," *Amer. J. Occup. Ther.*, vol. 39, no. 6, pp. 386–391, Jun. 1985.

[14] C. M. Light, P. H. Chappell, and P. J. Kyberd, "Establishing a standardized clinical assessment tool of pathologic and prosthetic hand function: Normative data, reliability, and validity," *Arch. Phys. Med. Rehabil.*, vol. 83, no. 6, pp. 776–783, 2002.

[15] H. Y. N. Lindner, J. M. Linacre, and L. M. N. Hermansson, "Assessment of capacity for myoelectric control: Evaluation of construct and rating scale," *J. Rehabil. Med.*, vol. 41, no. 6, pp. 467–474, 2009.

[16] L. A. Miller and S. Swanson, "Summary and recommendations of the academy's state of the science conference on upper limb prosthetic outcome measures," *J. Prosthetics Orthot.*, vol. 21, no. 9, pp. P83–P89, Oct. 2009.

[17] T. A. Kuiken *et al.*, "Targeted muscle reinnervation for real-time myoelectric control of multifunction artificial arms," *J. Amer. Med. Assoc.*, vol. 301, no. 6, pp. 619–628, Feb. 2009.

[18] A. M. Simon, L. J. Hargrove, B. A. Lock, and T. A. Kuiken, "Target achievement control test: Evaluating real-time myoelectric pattern-recognition control of multifunctional upper-limb prostheses," *J. Rehabil. Res. Develop.*, vol. 48, no. 6, pp. 619–627, 2011.

[19] E. J. Scheme and K. B. Englehart, "Validation of a selective ensemble-based classification scheme for myoelectric control using a three dimensional Fitts' law test," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 21, no. 4, pp. 616–623, Jul. 2013.

[20] P. M. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement," *J. Exp. Psychol.*, vol. 47, no. 6, pp. 381–391, 1954.

[21] X. Zhang and I. S. MacKenzie, "Evaluating eye tracking with ISO 9241—Part 9," in *Proc. Int. Conf. Hum.-Comput. Interact. (HCI)*, 2007, pp. 779–788.

[22] J. Park, W. Bae, H. Kim, and S. Park, "EMG—Force correlation considering Fitts' law," in *Proc. IEEE Int. Conf. Multisensor Fusion Integr. Intell. Syst.*, Aug. 2008, pp. 644–649.

[23] C. Choi, S. Micera, J. Carpaneto, and J. Kim, "Development and quantitative performance evaluation of a noninvasive EMG computer interface," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 1, pp. 188–191, Jan. 2009.

[24] M. R. Williams and R. F. Kirsch, "Evaluation of head orientation and neck muscle EMG signals as command inputs to a human–computer interface for individuals with high tetraplegia," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 16, no. 5, pp. 485–496, Oct. 2008.

[25] E. Mastinu, M. Ortiz-Catalan, and B. Håkansson, "Analog front-ends comparison in the way of a portable, low-power and low-cost EMG controller based on pattern recognition," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 2111–2114.

[26] L. J. Hargrove, E. J. Scheme, K. B. Englehart, and B. S. Hudgins, "Multiple binary classifications via linear discriminant analysis for improved controllability of a powered prosthesis," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 18, no. 1, pp. 49–57, Feb. 2010.

[27] M. Ortiz-Catalan, B. Håkansson, and R. Brånemark, "Real-time and simultaneous control of artificial limbs based on pattern recognition algorithms," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 4, pp. 756–764, Jul. 2014.

[28] R. W. Soukoreff and I. S. MacKenzie, "Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law Res. HCI," *Int. J. Hum. Comput. Stud.*, vol. 61, no. 6, pp. 751–789, 2004.

[29] L. van Dijk, C. K. van der Sluis, H. W. van Dijk, and R. M. Bongers, "Learning an EMG controlled game: Task-specific adaptations and transfer," *PLoS ONE*, vol. 11, no. 8, p. e0160817, 2016.

[30] F. Anderson and W. F. Bischof, "Augmented reality improves myoelectric prosthesis training," *Int. J. Disabil. Hum. Develop.*, vol. 13, no. 3, pp. 349–354, 2014.

JACOB GUSMAN received the B.Sc. degree in biomedical engineering from the University of Rochester, Rochester, NY, USA, in 2016. He is currently pursuing the Ph.D. degree in neural engineering with Brown University, Providence, RI, USA.

Between 2012 and 2016, he was a Student Intern with the Draper Laboratory, Cambridge, MA, USA, and in 2015, he was a Research Assistant with the Carney Laboratory, University of Rochester. In 2016, he received the Whitaker International Fellowship, whereby he spent a year with the Biomechatronics and Neurorehabiliation Laboratory, Chalmers University of Technology, Gothenburg, Sweden. His research interests include biological signal processing, prosthetic control, and brain-machine interfaces.

ENZO MASTINU (S'15) received the M.Sc. degree (Hons.) in electronic engineering from the Università degli Studi di Cagliari, Italy, in 2014. He is currently pursuing the Ph.D. degree with the Biomechatronics and Neurorehabilitation Laboratory (@ChalmersBNL), Department of Electrical Engineering, Chalmers University of Technology and Integrum AB, Gothenburg, Sweden. During and after the M.Sc. degree, he was an HW/FW Designer for a spin-off company and a Post-Graduate Research Scholar, respectively. His research interests include embedded systems for prosthetic control, mainly regarding upper limb amputees.

MAX ORTIZ CATALÁN (M'14) was born in Toluca, Mexico, in 1982. He received the Electronic Engineering degree from ITESM Campus, Toluca, in 2005, the M.Sc. degree in complex adaptive systems in 2009, and the Ph.D. degree in biomedical engineering from the Chalmers University of Technology (CTH), Gothenburg, Sweden, in 2014.

He is currently an Associate Professor with CTH, where he founded the Biomechatronics and Neurorehabilitation Laboratory (@ChalmersBNL). He also serves as the Research Director with Integrum AB, Sweden. His research interests include bioelectric signals acquisition electronics (analog and digital), signal processing and artificial intelligence algorithms for pattern recognition and control, neuromuscular interfaces, bone-anchored prostheses and osseointegration, virtual and augmented reality for neuromuscular rehabilitation, and the treatment of phantom limb pain.

Dr. Ortiz Catalán received the European Youth Award in 2014, the Delsys Prize in 2016, the SER Prize in 2017, and the Brian & Joyce Blatchford Award in 2017.